Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications 119 (2009) 897–923

stochastic processes and their applications

www.elsevier.com/locate/spa

Subgeometric rates of convergence of f-ergodic strong Markov processes

Randal Douc[a], Gersende Fort[b], Arnaud Guillin[c,*]

[a] *CMAP, Ecole Polytechnique, 91128 Palaiseau Cedex, France*
[b] *CNRS/LTCI, 46 rue Barrault, 75634 Paris Cedex 13, France*
[c] *Ecole Centrale Marseille and LATP, CNRS UMR 6632, 39 rue Joliot Curie, 13453 Marseille Cedex 13, France*

Received 20 March 2007; received in revised form 12 December 2007; accepted 28 March 2008
Available online 18 April 2008

Abstract

We provide a condition in terms of a supermartingale property for a functional of the Markov process, which implies (a) f-ergodicity of strong Markov processes at a subgeometric rate, and (b) a moderate deviation principle for an integral (bounded) functional. An equivalent condition in terms of a drift inequality on the extended generator is also given. Results related to (f, r)-regularity of the process, of some skeleton chains and of the resolvent chain are also derived. Applications to specific processes are considered, including elliptic stochastic differential equations, Langevin diffusions, hypoelliptic stochastic damping Hamiltonian systems and storage models.

MSC: primary 60J25; 37A25; secondary 60F10; 60J35; 60J60

Keywords: Subgeometric ergodicity; Regularity; Foster's criterion; Resolvent; Moderate deviations; Langevin diffusions; Hypoelliptic diffusions; Storage models

1. Introduction

In the present paper, we study the recurrence of continuous-time Markov processes. More precisely, we provide a criterion that yields a precise control of a subgeometric moment of the return-time to a test-set. The obtained result permits further quantitative analysis of

* Corresponding author.
E-mail addresses: douc@cmapx.polytechnique.fr (R. Douc), gfort@tsi.enst.fr (G. Fort), guillin@cmi.univ-mrs.fr, arnaud.guillin@ec-marseille.fr (A. Guillin).

0304-4149/$ - see front matter
doi:10.1016/j.spa.2008.03.007

characteristics such as the rate of convergence to the stationary state, a moderate deviation principle and the regularity of the process.

The stability and ergodic theory of continuous-time Markov processes have a large literature which is mainly devoted to the geometric case (also referred to as the exponential case). Meyn and Tweedie developed stability concepts for continuous-time Markov processes as well as simple criteria for non-explosivity, non-evanescence, Harris-recurrence, positive Harris-recurrence, ergodicity and geometric ergodicity [20,22,23]. Of particular importance in actually applying these concepts is the existence of verifiable conditions. In the discrete-time context, development of Foster–Lyapunov-type conditions on the transition kernel has provided such criteria (e.g. [21]). In the continuous-time context, Foster–Lyapunov inequalities applied to the generator of the process play the same role. These criteria were successfully applied to the study of the solution to stochastic differential equations (see [16] and more recently, [14] and references therein). Results relative to rates of convergence slower than geometric are not so well established. In [31,19] (resp. [33]), polynomial and subexponential ergodicity of stochastic differential equations (resp. subexponential ergodicity of queuing models) are addressed, but these results are quite model specific. Fort and Roberts [11] are, to our best knowledge, the first to study the subgeometric ergodicity of general strong Markov processes. Their conditions are in terms of subgeometric moment of the return-time to a test-set. Fort and Roberts derive nested drift inequalities on the generator of the process that makes the result one of practical interest in the polynomial case.

One of the applications of the condition we derive in the present paper makes the Fort–Roberts theory applicable for more general subgeometric rates such as the logarithmic or the subexponential ones. It also provides criteria for the (f, r)-regularity of a process, a characteristic which is an extension of the regularity concept [22]. We obtain theoretical results that are analogous to those in the discrete-time case [30]. We then relate our condition to a criterion based on the generator of the process. This criterion is the natural analogue of the Foster–Lyapunov condition for the geometric case; it also provides a single drift condition that generates the set of nested drift conditions by Fort–Roberts [11] for the polynomial case. Furthermore, it is analogous to the discrete-time version recently proposed by Douc–Fort–Moulines–Soulier [5].

In the literature, one approach for the theory of continuous-time Markov process is through the use of associated discrete-time chains: the resolvent chains and/or a skeleton chain. We discuss how our condition is related to a subgeometric drift inequality for these discrete-time Markov chains. As a consequence, we state new limit theorems such as moderate deviations for integrals of bounded functionals, thus weakening the conditions derived in Guillin–Wu [15,32].

Our conditions are then successfully applied to various non-trivial models: (a) we first consider elliptic stochastic differential equations for which conditions on the drift function enable us to generalize results by Veretennikov [31], Ganidis–Roynette–Simonot [12] or Malyshkin [19] (see also Pardoux–Veretennikov [26] for a study of the regularity of the solution of the Poisson equation under this drift condition); (b) we then study a "cold" Langevin tempered diffusion when the invariant target distribution is subexponential in the tails. This model is particularly useful in the Markov Chain Monte Carlo method. Different regimes of ergodicity (polynomial, subexponential or exponential) depending on the coldness of the diffusion term are exhibited, the different regimes are then characterized by the invariant target distribution. This study generalizes the Fort–Roberts results, which consider the case when the target density is polynomial in the tails [11]; (c) we also give a toy hypoelliptic example, namely a stochastic damping Hamiltonian system, in the case when it cannot be geometrically ergodic. This model is shown to be polynomially ergodic (see Wu [32] for the exponential case); (d) we finally consider a simple compound Poisson-process driven Ornstein–Uhlenbeck process (relevant for recent studies in financial econometrics) with a heavy tailed jump. It is shown to be subgeometrically ergodic.

Our approach may be considered as a probabilistic one. There are another ways to get subexponential rates of convergence (in total variation norm, in L^2 or in entropy) such as those based on spectral techniques (as in [12]), or on functional inequalities (weak Poincaré inequalities [28] or weak logarithmic Sobolev inequalities [2]). These results are however not easy to compare to ours and we postpone a comparative utilization of these approaches to further research.

Let us finally present the organization of the paper. Section 2 recalls basic definitions on Markov processes. The main results are given in Section 3: we provide a condition in terms of a supermartingale property for a functional of the Markov process, which is shown to imply (f, r)-ergodicity and moderate deviation principle for integral (bounded) functional. An equivalent formulation in terms of a drift inequality on the extended generator is also given. In Section 4, we consider basic properties characterizing (f, r)-ergodic Markov processes, such as the existence of suitably regular sets for the process or for some associated discrete-time Markov chains like the skeleton chains or the resolvent. All the proofs are given in the Appendix. Section 5 is devoted to the examples.

2. Definitions

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, (X_t)_{t\geq 0}, (\mathbb{P}_x)_{x\in\mathsf{X}})$ be a Markov family on a locally compact and separable metric space X endowed with its Borel σ-field $\mathcal{B}(\mathsf{X}) : (\Omega, \mathcal{F})$ is a measurable space, $(X_t)_{t\geq 0}$ be a Markov process with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$ and $\mathbb{P}_x$ (resp. $\mathbb{E}_x$) denote the canonical probability (resp. expectation) associated to the Markov process with initial distribution the point mass at x. Throughout this paper, the process is assumed to be a time-homogeneous strong Markov process with cad-lag paths, and we denote by $(P_t)_{t\geq 0}$ the associated semi-group on $(\mathsf{X}, \mathcal{B}(\mathsf{X}))$.

Let Λ_0 denote the class of the measurable and non-decreasing functions $r : [0, +\infty) \to [2, +\infty)$ such that $\log r(t)/t \downarrow 0$ as $t \to +\infty$. Let Λ denote the class of positive measurable functions $\bar{r}$, such that for some $r \in \Lambda_0$,

$$0 < \liminf_t \frac{\bar{r}(t)}{r(t)} \leq \limsup_t \frac{\bar{r}(t)}{r(t)} < \infty.$$

Λ is the class of the subgeometric rate functions and examples of functions $\bar{r} \in \Lambda$ are

$$\bar{r}(t) = t^{\alpha} (\log t)^{\beta} \exp(\gamma t^{\delta})$$

for $0 < \delta < 1$ and either $\gamma > 0$, or $\gamma = 0$ and $\alpha > 0$, or $\gamma = \alpha = 0$ and $\beta \geq 0$. We are ultimately interested in conditions implying that for all $x \in \mathsf{X}$

$$\lim_{t\to+\infty} r(t) \| P^t(x, \cdot) - \pi(\cdot) \|_f = 0, \tag{2.1}$$

where $r \in \Lambda$, π is the (unique) invariant distribution of the process i.e. $\pi P^t = \pi$ for all $t \geq 0$, and for a signed measure μ, $\|\mu\|_f = \sup_{|g|\leq f} |\mu(g)|$ where $f : \mathsf{X} \to [1, \infty)$ is a measurable function. When f is the constant function $\mathbf{1}$, the f-norm is nothing more than the total variation norm.

To attain that goal, we will need different notions of regularity and stability of continuous-time Markov processes and we briefly recall some basic definitions. The process is ϕ-irreducible for some σ-finite measure ϕ on $\mathcal{B}(\mathsf{X})$ if $\phi(A) > 0$ implies $\mathbb{E}_x\left[\int_0^\infty \mathbf{1}_A(X_s)\,\mathrm{d}s\right] > 0$ for all $x \in \mathsf{X}$. A ϕ-irreducible process possesses a maximal irreducibility measure ψ such that ϕ is absolutely continuous with respect to ψ for any other irreducibility measure ϕ [24]. Maximal irreducibility measures are not unique and are equivalent. A set $A \in \mathcal{B}(\mathsf{X})$ such that $\psi(A) > 0$ for some maximal irreducibility measure is said to be accessible; and full if $\psi(A^c) = 0$. A non-empty measurable set C is ν_a-petite (or simply petite) if there exist a probability measure a on the Borel σ-field of $[0, +\infty)$ and a non-trivial σ-finite measure ν_a on $\mathcal{B}(\mathsf{X})$ such that

$$\forall x \in C, \quad \int_0^{+\infty} P^t(x, \cdot)\, a(\mathrm{d}t) \geq \nu_a(\cdot).$$

For a ψ-irreducible process, an accessible closed petite set always exists [20]. A process is Harris-recurrent if, for some σ-finite measure μ, $\mu(A) > 0$ implies that the event $\{\int_0^\infty \mathbf{1}_A(X_s)\mathrm{d}s = \infty\}$ holds $\mathbb{P}_x$-a.s. for all $x \in \mathsf{X}$. Harris-recurrence trivially implies ϕ-irreducibility. A Harris-recurrent right process possesses an invariant measure π [13]; if π is an invariant probability distribution, the process is called positive Harris-recurrent. A ϕ-irreducible process is aperiodic if there exist an accessible ν_{δ_m}-petite set C and t_0 such that for all $x \in C, t \geq t_0$, $P^t(x, C) > 0$. A sufficient condition for a positive Harris-recurrent process to be aperiodic is the existence of some ϕ-irreducible skeleton chain [22]; recall that a skeleton P^m $(m > 0)$ is said to be ϕ-irreducible if there exists a σ-finite measure μ such that $\mu(A) > 0$ implies $\forall x \in \mathsf{X}, \exists k \in \mathbb{N}, P^{km}(x, A) > 0$ [21].

A ψ-irreducible and aperiodic Markov process that verifies (2.1) is said to be f-ergodic at a subgeometric rate (or simply f-ergodic when $r = \mathbf{1}$). When r is of the form $r(t) = \kappa^t$ for some $\kappa > 1$, the process is said to be f-ergodic at a geometric rate. In the literature, criteria for the stability of Markov processes, when stability is couched in terms of Harris-recurrence, positive Harris-recurrence, f-ergodicity, within this latter case, a mention of the rate of convergence, are expressed in terms of hitting times of some closed petite set. For any $\delta > 0$ and any closed set $C \in \mathcal{B}(\mathsf{X})$, let

$$\tau_C(\delta) = \inf\{t \geq \delta, X_t \in C\},$$

be the hitting time on C delayed by δ and define its (f, r)-modulated moment

$$G_C(x, f, r; \delta) = \mathbb{E}_x\left[\int_0^{\tau_C(\delta)} r(s)\; f(X_s)\,\mathrm{d}s\right],$$

where $f : \mathsf{X} \to [1, \infty)$ is a measurable function and $r : [0, +\infty) \to (0, +\infty)$ is a rate function. When $f = \mathbf{1}$ (resp. $r = \mathbf{1}$), this moment is simply called the r-modulated (resp. f-modulated) moment. Following discrete-time usage [21,30,17], we call a measurable set $C(f, r)$-regular if

$$\sup_{x \in C} G_B(x, f, r; \delta) < \infty,$$

for all $\delta > 0$ and all accessible sets B. Criteria for Harris-recurrence and positive Harris-recurrence can be found in [20, Theorems 1.1 and 1.2]; ergodicity and f-ergodicity are addressed in [22, Theorems 6.1 and 7.2]; criteria for geometric f-ergodicity at a geometric rate (resp. at a subgeometric rate) are provided by [7, Theorem 7.4] (resp. [11, Theorem 1]). A short review of these notions and results can be found in [11].

In many applications, these moments cannot be explicitly calculated; a second set of criteria based on the extended generator were thus derived for some of the stability properties above. We postpone to Section 3.3 a review of the existing conditions.

3. Main results

Let us consider the following drift condition towards a closed petite set C.

$\mathbf{D}(\mathbf{C}, \mathbf{V}, \phi, \mathbf{b})$: There exist a closed petite set C, a continuous function $V : \mathsf{X} \to [1, \infty)$, an increasing differentiable concave positive function $\phi : [1, \infty) \to (0, \infty)$ and a constant $b < \infty$ such that for any $s \geq 0$, $x \in \mathsf{X}$,

$$\mathbb{E}_x [V(X_s)] + \mathbb{E}_x \left[\int_0^s \phi \circ V(X_u) \mathrm{d}u \right] \leq V(x) + b\, \mathbb{E}_x \left[\int_0^s \mathbf{1}_C(X_u) \mathrm{d}u \right]. \tag{3.1}$$

Note that (3.1) is equivalent to the condition that the functional

$$s \mapsto V(X_s) - V(X_0) + \int_0^s \phi \circ V(X_u) \mathrm{d}u - b \int_0^s \mathbf{1}_C(X_u) \mathrm{d}u \tag{3.2}$$

is, for all $x \in \mathsf{X}$, a $\mathbb{P}_x$-supermartingale with respect to the filtration $(\mathcal{F}_t)_{t \geq 0}$.

We prove that this condition is related to (i) the f-ergodicity at a subgeometric rate of the Markov process, (ii) a Moderate Deviation Principle for additive integral of bounded functional of the Markov process, (iii) a drift inequality on the extended generator.

3.1. f-ergodicity at a subgeometric rate

The first result concerns the existence of an invariant probability distribution π and shows that the drift condition $\mathbf{D}(\mathbf{C}, \mathbf{V}, \phi, \mathbf{b})$ provides a simple tool when identifying the set of the π-integrable functions.

Proposition 3.1. *Assume* $\mathbf{D}(\mathbf{C}, \mathbf{V}, \phi, \mathbf{b})$ *and* $\sup_C V < \infty$. *Then the process is positive Harris-recurrent with an invariant probability measure* π *such that* $\pi(\phi \circ V) < \infty$.

Proposition 3.1 results from [20, Theorems 1.1 and 1.2] and Theorem 4.1(i). It is known that positive Harris-recurrence does not necessarily imply ergodicity and aperiodicity is required [22, Proposition 6.1]; similar conditions are required in the discrete-time case [21]. In the present case, more information than positive Harris-recurrence is available and thus, f-ergodicity at a subgeometric rate can be established.

Set

$$H_\phi(u) = \int_1^u \frac{\mathrm{d}s}{\phi(s)}, \quad u \geq 1.$$

Theorem 3.2 states that the Markov process converges in f-norm to the invariant probability measure π, for a wide family of functions $1 \leq f \leq f_*$ and a wide family of rate functions $r \leq r_*$ where

$$f_* = \phi \circ V, \qquad r_*(s) = \phi \circ H_\phi^{-1}(s).$$

To attain that goal, we introduce the pairs of Young's functions (H_1, H_2) that – among other properties – satisfy the property

$$x\, y \leq H_1(x) + H_2(y), \quad \forall x, y \geq 0, \tag{3.3}$$

and are invertible (see e.g [18, Chapter 1]). Examples of pairs (H_1, H_2) are given in [5,11] while a general construction can be found in [18, Chapter 1]. Let $\mathcal{I}$ be the pairs of inverse Young's functions augmented with the pairs (Id, $\mathbf{1}$) and ($\mathbf{1}$, Id): $(\Psi_1, \Psi_2) \in \mathcal{I}$ iff $(\Psi_1^{-1}, \Psi_2^{-1})$ is a pair of Young's functions. For a rate function $r \in \Lambda$, define $r^0(t) = \int_0^t r(s)\,\mathrm{d}s$, and, if r is differentiable, set $\partial r(t) = \frac{\mathrm{d}r(t)}{\mathrm{d}t}$.

Theorem 3.2. *Assume that:*

(i) *some skeleton chain is irreducible.*
(ii) *the condition* $\mathbf{D}(\mathbf{C}, \mathbf{V}, \phi, \mathbf{b})$ *holds with* C, V, ϕ *such that* $\sup_C V < \infty$ *and* $\lim_{+\infty} \phi' = 0$.

For any pair $\Psi = (\Psi_1, \Psi_2) \in \mathcal{I}$ *and any probability distribution* λ *satisfying* $\lambda(V) < \infty$,

$$\lim_{t\to+\infty} \{\Psi_1(r_*(t)) \vee 1\} \int_{\mathsf{X}} \lambda(\mathrm{d}x)\, \|P^t(x,\cdot) - \pi(\cdot)\|_{\Psi_2(f_*)\vee 1} = 0. \tag{3.4}$$

Furthermore, there exist finite constants $C_{\Psi,i}$ *such that for all* $t \geq 0$ *and all* $x \in \mathsf{X}$,

$$\{\Psi_1(r_*(t)) \vee 1\}\, \|P^t(x,\cdot) - \pi(\cdot)\|_{\Psi_2(f_*)\vee 1} \leq C_{\Psi,1}\, V(x), \tag{3.5}$$

$$\int_0^\infty \{\Psi_1(r_*(t)) \vee 1\}\, \|P^t(x,\cdot) - P^t(y,\cdot)\|_{\Psi_2(f_*)\vee 1}\, \mathrm{d}t \leq C_{\Psi,2}\, \{V(x) + V(y)\}; \tag{3.6}$$

and if $\partial[\Psi_1(r_*)] \in \Lambda$, *there exists a finite constant* $C_{\Psi,3}$ *such that for all* $t \geq 0$,

$$\int_0^\infty \{\partial[\Psi_1(r_*)](t) \vee 1\} \quad \|P^t(x,\cdot) - \pi(\cdot)\|_{\Psi_2(f_*)\vee 1}\, \mathrm{d}t \leq C_{\Psi,3}\, V(x). \tag{3.7}$$

Limit (3.4) is a direct application of [11, Theorem 1] while (3.5)–(3.7) are, to our best knowledge, new results for the continuous-time Markov process theory. The proof is detailed in Appendix. Observe that under the stated assumptions, $\Psi_1 \circ r_* \in \Lambda$ (see Lemma 4.3).

As said in [11], Eq. (3.4) shows that the rate of convergence and the norm in which convergence occurs have to be balanced: for two pairs (Ψ_1, Ψ_2) and (Ψ_1', Ψ_2') in $\mathcal{I}$, if $\Psi_1(x) \leq \Psi_1'(x)$ for all large x, then $\Psi_2(y) \geq \Psi_2'(y)$ for all large y [18, Theorem 1.2.1]. Hence, the stronger the norm, the weaker the rate and conversely. The maximal rate of convergence is achieved with the total variation norm ($\Psi_2 \circ f_* = \mathbf{1}$) and the minimal one ($\Psi_1 \circ r_* = \mathbf{1}$) is achieved with the f_*-norm. Hence, the drift condition $\mathbf{D}(\mathbf{C}, \mathbf{V}, \phi, \mathbf{b})$ directly provides information on the rate of convergence and the norm of convergence: the largest rate of convergence $r_* = \phi \circ H_\phi^{-1}$ is given by the concave function ϕ and the largest norm of convergence $\|\cdot\|_{f_*}$ is given by the pair (ϕ, V). Eqs. (3.5)–(3.7) are, to our best knowledge, the first results that address the dependence upon the initial point in the ergodic behavior. When applied to discrete-time Markov chains, (3.5) to (3.7) coincide with resp. [30, Theorems 2.1, 4.1, 4.2] (the dependence upon x can be read from the proof of these theorems; the details are also provided in [9, Chapter 3]). These results for the discrete-time case and the definition of the set $\mathcal{S}_\psi$ in [11, Theorem 1] suggest that in (3.5), the minimal dependence on the starting value x is of the form $G_C(x, \Psi_2(f_*), \Psi_1(r_*); \delta)$. Similar expressions can be predicted for (3.6) and (3.7). The proof of this assertion and the explicit construction of the constants $C_{\Psi,i}$ in terms of the quantities appearing in the assumptions are beyond the scope of this paper. The work on explicit control of subgeometric ergodicity for strong Markov processes is currently in progress.

In the examples given in Section 5, we will see that the pair (ϕ, V) that solves $\mathbf{D}(\mathbf{C}, \mathbf{V}, \phi, \mathbf{b})$ is not unique, that is, the drift condition only provides an upper bound of the true rate of convergence (see e.g. Section 5.1). Nevertheless, in many applications, we are able to prove that the true rate belongs to the exhibited class of rate functions (see e.g. Section 5.2).

3.2. Skeleton chain and moderate deviations

We consider here an important field of application for this subgeometric rate, namely moderate deviations for bounded additive functionals of Markov process. Moderate deviations are concerned with the asymptotic for centered g with respect to π and for $0 \le t \le T$ of

$$S_t^\epsilon = \frac{1}{\sqrt{\epsilon}h(\epsilon)} \int_0^t g(X_{s/\epsilon})\mathrm{d}s$$

where as ϵ tends to 0, $h(\epsilon) \to \infty$ but $\sqrt{\epsilon}h(\epsilon) \to 0$, namely a regime between the large deviations and the central limit theorem. Let $C_0([0,1], \mathbb{R}^n)$ be the space of continuous functions from [0, 1] to $\mathbb{R}^n$ starting from 0 equipped with the supremum norm topology. We may then state:

Theorem 3.3. *Assume that* $\mathbf{D}(\mathbf{C}, \mathbf{V}, \phi, \mathbf{b})$ *holds with* $\sup_C V < \infty$, *and some skeleton chain is* ψ*-irreducible.*

(i) *For all* $m > 0$, *there exist a function* $W : \mathsf{X} \to [\phi(1), \infty)$, *an accessible petite set* $\tilde{C}$ *for the skeleton* P^m *and a positive constant* b' *such that* $\sup_{\tilde{C}} W$ *is finite, and on* X,

$$P^m W \le W - \phi \circ W + b' \mathbf{1}_{\tilde{C}}, \quad \text{and} \quad \phi \circ V \le W \le \kappa V.$$

(ii) *Suppose that there exists a function* h *such that* $h(\epsilon) \to \infty$ *but* $\sqrt{\epsilon}h(\epsilon) \to 0$ *and for all* $a > 0$,

$$\lim_{\epsilon \to 0} \frac{1}{h^2(\epsilon)} \log\left(\epsilon H_\phi^{-1}\left(\frac{a\, h(\epsilon)}{\sqrt{\epsilon}}\right)\right) = +\infty. \tag{3.8}$$

Suppose that $\pi(V) < \infty$ *and set*

$$\sigma^2(\langle g, \zeta\rangle) = \lim_{n\to\infty} \frac{1}{n}\mathbb{E}_\pi \left(\int_0^n \langle g, \zeta\rangle(X_s)\mathrm{d}s\right)^2 = 2\int_{\mathsf{X}} \langle g, \zeta\rangle \int_0^\infty P^t \langle g, \zeta\rangle \mathrm{d}t\mathrm{d}\pi.$$

Then, for any initial distribution μ *such that* $\mu(V) < +\infty$, $\mathbb{P}_\mu\left(S_\cdot^\epsilon \in \cdot\right)$ *satisfies a moderate deviation principle in* $C_0([0,1], \mathbb{R}^n)$ *with speed* $\frac{1}{h^2(\epsilon)}$ *and rate function* I_g^h *given by*

$$I_g^h(\gamma, \sigma) = \frac{1}{2}\int_0^1 \sup_{\zeta\in\mathbb{R}^n} \left\{\langle \dot{\gamma}(t), \zeta\rangle - \frac{1}{2}\sigma^2(\langle g, \zeta\rangle)\right\} \mathrm{d}t \tag{3.9}$$

if $d\gamma(t) = \dot{\gamma}(t)\mathrm{d}t$ *and* $\gamma(0) = 0$; *and* $I_g^h(\gamma, \sigma) = +\infty$ *otherwise.*

The proof is in Appendix. Note that $\sigma^2(\langle g, \zeta\rangle)$ is the usual variance of the Central Limit Theorem. To the authors' knowledge, this moderate deviation result (even for bounded function) is the first one for Markov processes which are not exponentially ergodic. It extends then results of Guillin [15, Th 1.] or Wu [32, Th. 2.7] in the subexponential setting. As expected, all ranges of speed are not allowed for such a theorem but are limited by the ergodicity of the process as stated by condition (3.8) which is however easy to check (we refer the reader to Douc–Guillin–Moulines [6, Sect. 4] for a complete discussion on this interplay). For example, suppose that you are in the subexponential regime, say $\phi(v) = v\log(v)^{2p/(1-p)}$ for some $0 < p < 1$ (see Section 5.1 for assumptions ensuring that a diffusion process satisfies such a condition), then (3.8) implies that the speed should satisfy $1 \ll h(\epsilon) \ll \epsilon^{(p-1)/2(1+3p)}$. Note that as a direct application a functional LIL may be easily obtained under a quite optimal ergodicity condition.

3.3. Generator and drift inequality (3.1)

The drift condition $\mathbf{D}(\mathbf{C}, \mathbf{V}, \phi, \mathbf{b})$ may be not easy to derive since it is couched in terms of the process itself. The main goal of this section is to provide a sufficient condition similar to the usual form of conditions adopted in an earlier paper to address different classes of stability: namely a condition based on the extended generator [4, Def. 1.15.15].

Let $\mathcal{D}(\mathcal{A})$ denote the set of measurable functions $f : \mathsf{X} \to \mathbb{R}$ such that: there exists a measurable function $h : \mathsf{X} \to \mathbb{R}$ such that the function $t \mapsto h(X_t)$ is integrable $\mathbb{P}_x$-a.s. for each $x \in \mathsf{X}$ and the process

$$t \mapsto f(X_t) - f(X_0) - \int_0^t h(X_s)\mathrm{d}s \tag{3.10}$$

is a $\mathbb{P}_x$-local martingale for all x. Then we write $h = \mathcal{A}f$, and f is said to be in the domain of the extended generator $(\mathcal{A}, \mathcal{D}(\mathcal{A}))$ of the process X. The condition (3.1) looks like a Dynkin formula. This is the reason why we want it to hold as widely as possible, thus justifying the interest in the extended generator concept.

Theorem 3.4. *Let $V : \mathsf{X} \to [1, \infty)$ with $V \in \mathcal{D}(\mathcal{A})$ be a continuous function and $\phi : [1, \infty) \to (0, \infty)$ be an increasing differentiable concave positive function.*

(i) *If there exist a closed petite set C and a constant $b < \infty$ such that for all $x \in \mathsf{X}$,*

$$\mathcal{A}V(x) \le -\phi \circ V(x) + b\mathbf{1}_C(x), \tag{3.11}$$

then $\mathbf{D}(\mathbf{C}, \mathbf{V}, \phi, \mathbf{b})$ holds.

(ii) *If $\mathbf{D}(\mathbf{C}, \mathbf{V}, \phi, \mathbf{b})$ holds for some compact set C and $\mathcal{A}V$ is continuous then (3.11) holds.*

The proof is in Appendix. The extended generator is less restrictive than the infinitesimal generator $\tilde{\mathcal{A}}$: if f is in the domain of $\tilde{\mathcal{A}}$, then the process (3.10) is a martingale and f is in the domain of $\mathcal{A}$ (see e.g. [4, Proposition 1.14.13]). In particular, it is often quite difficult to characterize the domain of $\tilde{\mathcal{A}}$ but there may be (and are, in the applications of Section 5) easily checked sufficient conditions for membership of $\mathcal{D}(\mathcal{A})$.

This drift condition is naturally a part of the existing literature, that addresses criteria for non-explosivity, recurrence, f-ergodicity, polynomial ergodicity, geometric and uniform ergodicity (see [23, Conditions (CD0) to (CD3)]; see also [7]). For example, (3.11) is an extension of

$$f\text{-ergodicity:}\quad \mathcal{A}V(x) \le -cf(x) + b\mathbf{1}_C(x), \tag{3.12}$$

$$\text{geometric ergodicity:}\quad \mathcal{A}V(x) \le -cV(x) + b\mathbf{1}_C(x) \tag{3.13}$$

where $V \ge 1$ and $f \ge 0$ are measurable functions, C is a closed petite set such that $\sup_C V < \infty$, b and c are positive constants; these two conditions resp. address the f-ergodicity and the geometric ergodicity. These criteria are similar to some conditions provided by [16] for the stability of stochastic differential equations. The drift inequality (3.13) is the limit of our approach, since it corresponds to (3.11) with $\phi(v) \propto v$.

In a recent work, Fort and Roberts [11] considered a family of drift conditions that imply f-ergodicity at a polynomial rate: namely, there exist $0 < \alpha < 1$, $b > 0$ such that for all $\alpha \le \eta \le 1$, there exists $c_\eta > 0$ such that

$$\mathcal{A}V^\eta(x) \le -c_\eta V^{\eta-\alpha}(x) + b1_C(x). \tag{3.14}$$

The comparison of the Fort–Roberts nested drift conditions (3.14) and our single drift condition can be more explicit when $V \in \mathcal{D}(\mathcal{A})$ and the process (3.10) is a $\mathbb{P}_x$-martingale for all x. In that case, it is easily seen that the single drift condition implies the nested drift conditions.

4. Modulated moments for the process, and associated discrete-time Markov chains

We present here byproduct results, that are related to the notion of (f, r)-regularity. The main result of Section 4.1 is Theorem 4.1 that states that this drift condition allows the calculation of an upper bound for some r-modulated moment where $r \in \Lambda$, and for some f-modulated moment, $f \geq 1$. Using interpolating inequalities, we obtain (f, r)-modulated moments for a wide family of pairs (f, r). Section 4.2 is devoted to (f, r)-regularity: the main result of this section is Proposition 4.7 that identifies (f, r)-regular sets from the condition $\mathbf{D}(\mathbf{C}, \mathbf{V}, \phi, \mathbf{b})$. We present in Section 4.4 the interplay between a drift condition on the resolvent kernel and the drift condition $\mathbf{D}(\mathbf{C}, \mathbf{V}, \phi, \mathbf{b})$.

All the proofs are given in Appendix.

4.1. Modulated moments for the process

Theorem 4.1. *Assume* $\mathbf{D}(\mathbf{C}, \mathbf{V}, \phi, \mathbf{b})$.

(i) *For all* $x \in \mathsf{X}$ *and* $\delta > 0$, $\mathbb{E}_x\left[\int_0^{\tau_C(\delta)} \phi \circ V(X_s)\, \mathrm{d}s\right] \leq V(x) - 1 + b\delta$.

(ii) *For all* $x \in \mathsf{X}$ *and* $\delta > 0$, $\mathbb{E}_x\left[\int_0^{\tau_C(\delta)} \phi \circ H_\phi^{-1}(s)\, \mathrm{d}s\right] \leq V(x) - 1 + \frac{b}{\phi(1)} \int_0^\delta \phi \circ H_\phi^{-1}(s)\mathrm{d}s$.

The proof of Theorem 4.1 does not require C to be petite. Nevertheless, this petiteness property will be crucial in all the following results: we will see that this assumption allows the extension of the above controls to those of modulated moments $\tau_B(\delta)$ for any accessible set B. Theorem 4.1 gives the largest f-modulated and r-modulated moments of $\tau_C(\delta)$ that can be deduced from $\mathbf{D}(\mathbf{C}, \mathbf{V}, \phi, \mathbf{b})$. Interpolated (f, r)-modulated moments of $\tau_C(\delta)$ can easily be obtained for a wide family of functions $1 \leq f \leq f_*$ (and, equivalently, a wide family of rate functions $r(s) \leq r_*(s)$). To attain that goal, we follow the same lines as in [5,11]. Corollary 4.2 trivially results from Theorem 4.1 and Eq. (3.3).

Corollary 4.2. *Assume* $\mathbf{D}(\mathbf{C}, \mathbf{V}, \phi, \mathbf{b})$. *For any pairs* $(\Psi_1, \Psi_2) \in \mathcal{I}$ *and all* $\delta > 0$,

$$\mathbb{E}_x\left[\int_0^{\tau_C(\delta)} \Psi_1\left(r_*(s)\right)\, \Psi_2\left(f_*(X_s)\right)\, \mathrm{d}s\right] \leq 2(V(x) - 1) + b\int_0^\delta \left(1 + \frac{r_*(s)}{r_*(0)}\right)\mathrm{d}s.$$

If Ψ_1 strongly increases at infinity then Ψ_2 slowly increases: the rate r and the function f have to be balanced (see [18] and the comments in Section 4.1).

Theorem 4.1 and Corollary 4.2 thus provide a control of (f, r)-modulated moments; which is of great interest when $\Psi_1 \circ r_*$ is a subgeometric rate function. A sufficient condition is Lemma 4.3 (see [5, Lemmas 2.3 and 2.7])

Lemma 4.3. *If* $\lim_\infty \phi' = 0$, *then* $r_* \in \Lambda$ *and for all inverse Young functions* Ψ_1, $\Psi_1 \circ r_* \in \Lambda$.

Proposition 4.4. *Assume* $\mathbf{D}(\mathbf{C}, \mathbf{V}, \phi, \mathbf{b})$. *Then the process is* ψ*-irreducible. If* $\sup_C V < \infty$,

(i) *the level sets* $\{V \leq n\}$ *are petite.*

(ii) *there exists a closed accessible petite set* B *such that* $\mathbf{D}(\mathbf{B}, \mathbf{V}, \phi, \mathbf{b})$ *holds and* $\sup_B V < \infty$.

Therefore, when $\mathbf{D}(\mathbf{C}, \mathbf{V}, \phi, \mathbf{b})$ holds and $\sup_C V < \infty$, we can assume without loss of generality (w.l.g.) that C is accessible.

4.2. (f, r)-regularity of the process

The objective of this section is to identify regular sets from the drift condition $\mathbf{D}(\mathbf{C}, \mathbf{V}, \phi, \mathbf{b})$. Proposition 4.5 shows that the "self-regularity" of a closed petite set C actually implies (f, r)-regularity. This results extends [20, Proposition 4.1] (resp. [11, Proposition 22]) that addresses the case $r = \mathbf{1}$ (resp. $f = \mathbf{1}$). It also generalizes [11, Proposition 23] which concerns the case $r = \Psi_1(r_*)$ and $f = \Psi_2(f_*)$ for some pair $(\Psi_1, \Psi_2) \in \mathcal{I}$. This proposition is the counterpart in the subexponential setting of the result by Down et al. for the exponential case [7, Theorem 7.2].

Proposition 4.5. *Let $f : \mathsf{X} \to [1, \infty)$ be a measurable function and $r \in \Lambda$ be a subgeometric rate function. Assume that the process is ψ-irreducible and $\sup_C G_C(\cdot, f, r; \delta) < \infty$ for some (and thus any) $\delta > 0$ and some closed petite set C. Then,*

(i) *the set $\{x \in \mathsf{X}, G_C(x, f, r; \delta) < \infty\}$ is full,*
(ii) *for all accessible sets $B \in \mathcal{B}(\mathsf{X})$ and all $t \geq 0$, there exists a constant $c_{B,t} < \infty$ such that $G_B(\cdot, f, r; t) \leq c_{B,t}\, G_C(\cdot, f, r; \delta)$ so that C is (f, r)-regular.*

Combining Proposition 4.5(ii), Corollary 4.2 and Lemma 4.3 yields the following result

Proposition 4.6. *Assume that $\mathbf{D}(\mathbf{C}, \mathbf{V}, \phi, \mathbf{b})$ holds with C, V, ϕ such that $\sup_C V < \infty$ and $\lim_{+\infty} \phi' = 0$. Then for any pair $(\Psi_1, \Psi_2) \in \mathcal{I}$, any accessible set B and all $\delta > 0$, there exists a finite constant c such that*

$$\mathbb{E}_x \left[\int_0^{\tau_B(\delta)} \Psi_1\left(r_*(s)\right) \Psi_2\left(f_*(X_s)\right)\, \mathrm{d}s \right] \leq c\, V(x),$$

and any V-level set $\{x \in \mathsf{X}, V(x) \leq v\}$ is $(\Psi_2 \circ f_*, \Psi_1 \circ r_*)$-regular.

We now establish a general result that extends to continuous-time Markov processes, part of [30, Theorem 2.1] relative to the discrete-time Markov chain. In the case $r = \mathbf{1}$, some of these equivalences are proved in [20] for continuous-time strong Markov processes.

Proposition 4.7. *Let $f : \mathsf{X} \to [1, \infty)$ be a measurable function and $r \in \Lambda$ be a subgeometric rate function. Assume that the process is ψ-irreducible. The following conditions are equivalent.*

(i) *There exist a closed petite set C and $\delta > 0$ such that $\sup_C G_C(x, f, r; \delta) < \infty$.*
(ii) *There exists a (f, r)-regular closed set which is accessible.*
(iii) *There exists a full set $\mathcal{S}_\Psi$ which is the union of a countable number of (f, r)-regular sets.*

By Theorem 4.1, these equivalent conditions are verified provided $\mathbf{D}(\mathbf{C}, \mathbf{V}, \phi, \mathbf{b})$ holds, $\sup_C V < \infty$ and $\lim_\infty \phi' = 0$.

4.3. Modulated moments for an irreducible skeleton

Under mild additional conditions, the drift condition $\mathbf{D}$ also yields controls of modulated moments for irreducible skeleton chains. For all $m > 0$, let $T_{m,C}$ be the return-time to C of the skeleton chain P^m,

$$T_{m,C} = \inf\{k \geq 1, X_{mk} \in C\}.$$

Proposition 4.8. *Assume that* $\mathbf{D}(\mathbf{C}, \mathbf{V}, \phi, \mathbf{b})$ *holds with* $\sup_C V < \infty$*, and some skeleton chain is irreducible. For all* $m > 0$ *and any accessible set* B*, there exist constants* c_i, $1 \leq i \leq 4$*, such that for all* $x \in \mathsf{X}$,

$$\mathbb{E}_x \left[\sum_{k=0}^{T_{m,B}-1} \phi \circ V(X_{mk}) \right] \leq c_1 \, \mathbb{E}_x \left[\int_0^{T_{m,B}} \phi \circ V(X_{sm}) \, \mathrm{d}s \right] \leq c_2 \, V(x),$$

$$\mathbb{E}_x \left[\sum_{k=0}^{T_{m,B}-1} r_*(km) \right] \leq c_3 \mathbb{E}_x \left[\int_0^{mT_{m,B}} r_*(s) \, \mathrm{d}s \right] \leq c_4 \, V(x).$$

4.4. Resolvent and drift inequality (3.1)

One of the approaches for studying the stability and ergodic theory of continuous-time Markov processes consists in making use of the associated discrete-time resolvent chains. This allows taking profit of the analysis of discrete-time Markov chains which is quite well understood [24,21] and then transferring properties established in terms of the resolvent or "generalized resolvent" kernel (see for e.g. [20]) to the Markov process itself. More precisely, define, for $\beta > 0$, the resolvent kernel R_β by $R_\beta(x, A) = \int_0^\infty \beta \mathrm{e}^{-\beta t} P^t(x, A) \mathrm{d}t$ and consider the following drift condition associated to the resolvent kernel.

$\check{\mathbf{D}}(\mathbf{C}, \mathbf{V}, \phi, \mathbf{b}, \beta)$: There exist a petite set C, a function $V : \mathsf{X} \to [1, \infty)$, an increasing differentiable concave positive function $\phi : [1, \infty) \to (0, \infty)$ and a constant $b < \infty$ such that for any $x \in \mathsf{X}$,

$$R_\beta V(x) \leq V(x) - \phi \circ V(x) + b \mathbf{1}_C(x). \tag{4.1}$$

Theorem 4.9 ensures that drift conditions expressed in terms of the resolvent kernel or of the Markov process are essentially equivalent. This theorem parallels [7, Theorem 5.1] which has been established for exponentially ergodic Markov processes. The proof is given in Appendix.

Theorem 4.9. (i) *Assume* $\check{\mathbf{D}}(\mathbf{C}, \mathbf{V}, \phi, \mathbf{b}, \beta)$ *where* C *is a closed set and* $R_\beta V$ *is a continuous function. Then* $\mathbf{D}(\mathbf{C}, \mathbf{R}_\beta \mathbf{V}, \beta\phi, \beta\mathbf{b})$ *holds.*

(ii) *Assume* $\mathbf{D}(\mathbf{C}, \mathbf{V}, \phi, \mathbf{b})$ *with* $\sup_C V < \infty$*. Then, for all* $\beta > 0$*, there exist* $(\check{C}, W, \check{\phi}, \check{b})$ *such that* $\check{\mathbf{D}}(\check{\mathbf{C}}, \mathbf{W}, \check{\phi}, \check{\mathbf{b}}, \beta)$ *holds.* $\check{C}$*,* W *and* $\check{\phi}$ *are such that*

$$\sup_{\mathsf{X}} \left(W - V \left(1 + \phi'(1) \right) \right) < \infty, \qquad \sup_{\check{C}} W < \infty, \qquad \check{\phi}(t(1 + \phi'(1))) \sim_{+\infty} \phi(t).$$

Modulated moments for the resolvent can be deduced from (4.1) by applying the results by [5]. The details are omitted.

5. Examples

In this section, $\mathsf{X} = \mathbb{R}^n$. For a set A, A^c is its complement in $\mathbb{R}^n$. Vectors are intended as column vectors, $|x|$ and $\langle \cdot, \cdot \rangle$ denote respectively the Euclidean norm and the scalar product. For a matrix a, $|a| = \left(\sum_{i,j} a_{i,j}^2 \right)^{1/2}$, $\mathrm{Tr}(a)$ stands for the trace of the matrix and a' the matrix transpose. Id_n is the $n \times n$ identity matrix. If V is a twice continuously differentiable function w.r.t. $x \in \mathbb{R}^n$, ∂V (or $\partial_x V$ when confusion is possible) denotes its gradient, and $\partial^2 V$ its Hessian.

Four applications are considered: we analyze three different diffusions (a general elliptic diffusion on $\mathbb{R}^n$, a Langevin diffusion on $\mathbb{R}^n$ and an hypoelliptic diffusion) and a compound Poisson-process driven Ornstein–Uhlenbeck process. Queuing theory is another important field of application for our theory. We do not discuss here this field of applications. This will be done in a forthcoming paper, which will also include a comparison of our results to those by [3,33]. Techniques in Dai–Meyn [3] differ from ours since they are based on fluid limits. Concerning [33], our conditions are more general; indeed the authors assume that there exists a state x_0 such that whenever the Markov process hits x_0, it will sojourn there for a random time that is positive with probability 1, [33, Assumption 1.1]. This assumption makes their results unavailable for the applications we now consider.

5.1. *Elliptic diffusions on* $\mathbb{R}^n$

Consider the stochastic integral equation of the form

$$X_t = X_0 + \int_0^t b(X_s)\mathrm{d}s + \int_0^t \sigma(X_s)\mathrm{d}B_s, \tag{5.1}$$

where $X_t \in \mathbb{R}^n$, $b : \mathbb{R}^n \to \mathbb{R}^n$ and $\sigma : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ are measurable functions, and $\{B_t\}_t$ is an n-dimensional Brownian motion. Assume that $b : \mathbb{R}^n \to \mathbb{R}^n$ and $\sigma : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ are functions satisfying,

A1 σ is bounded and b and σ are locally Lipschitz: for any $l > 0$, there exists a finite constant c_l such that for all $|x| \le l$, $|y| \le l$, $|b(x) - b(y)| + |\sigma(x) - \sigma(y)| \le c_l|x - y|$.

Let $a(x) = \sigma(x)\sigma(x)'$ be the diffusion matrix. We assume that,

A2 (i) $a(x)$ is non-singular: the smallest eigenvalue of the diffusion matrix $a(x)$ is bounded away from zero in every bounded domain.
(ii) there exist $0 < p < 1$, $r > 0$ and M such that for all $|x| \ge M$, $\langle b(x), x\rangle \le -r|x|^{1-p}$.

This implies that $\Lambda = n^{-1}\sup_{x\in\mathbb{R}^n} \mathrm{Tr}(a(x))$ and $\lambda_+ = \sup_{x\neq 0}\langle a(x)\frac{x}{|x|}, \frac{x}{|x|}\rangle$ are finite. Moreover, A2(i) is equivalent to the condition $\det(\sigma(x)) \neq 0$ for all x. Finally the process is regular, sufficient conditions for regularity can be found in [16, Theorem 3.4.1.] and there exists a solution to (5.1), which is an almost surely continuous stochastic process and is unique up to equivalence. This solution is an homogeneous Markov process whose semi-group is Feller [16, Theorem 3.4.1]. Hence, it is strongly Markovian, as a right-continuous Markov process with Feller semi-group.

Proposition 5.1. *Under A1–A2, X possesses an unique invariant probability measure π. π is a maximal irreducibility measure and any skeleton P^m is irreducible. Furthermore, the compact sets are closed petite sets.*

The proof results from classical results (see for example [16] and [21, Proposition 6.2.8] for the petiteness of the compact sets) and is omitted. Define the operator L that acts on function $V : \mathbb{R}^n$, $x \mapsto V(x)$ that are twice continuously differentiable, by

$$LV(x) = \langle b(x), \partial V(x)\rangle + \frac{1}{2}\mathrm{Tr}\left(a(x)\ \partial^2 V(x)\right). \tag{5.2}$$

By standard properties of the stochastic integral, any twice continuously differentiable function $V : \mathbb{R}^n \to \mathbb{R}$ is in the domain of $\mathcal{A}$ and $LV(x) = \mathcal{A}V(x)$ for all $x \in \mathbb{R}^n$.

Proposition 5.2. *Assume A1–A2. Let $V : \mathbb{R}^n \to [1, +\infty)$ be a twice continuously differentiable function such that $V(x) = \exp(\iota\, |x|^m)$ outside a compact set, for some $0 < m < 1$ and $\iota > 0$. Then $\sup_{|x|\le M} \mathcal{A}V(x) < \infty$ and for all $|x| \ge M$,*

(i) *If $0 < m < 1 - p$, $\mathcal{A}V(x) \le -\iota^{\frac{1+p}{m}} mr\, [\ln V(x)]^{1-(\frac{1+p}{m})}\, V(x)\,(1 + o(1))$.*
(ii) *If $m = 1 - p$ and $\{r - (1/2)\lambda_+\iota(1 - p)\} > 0$,*

$$\mathcal{A}V(x) \le -\iota^{\frac{1+p}{1-p}}(1-p)\,\{r - (1/2)\lambda_+\iota(1-p)\}\,[\ln V(x)]^{-2\frac{p}{1-p}}\, V(x)\,(1 + o(1))\,.$$

Proof. $\sup_{\{x,|x|\le M\}} \mathcal{A}V < \infty$ and by the definition of $\mathcal{A}$, for all $|x| \ge M$,

$$\mathcal{A}V(x) \le -\iota m\left(r - (1/2)\lambda_+\iota m|x|^{p+m-1}\right)|x|^{m-1-p}V(x) + (1/2)\iota m n\Lambda|x|^{m-2}V(x). \quad \square$$

Theorem 5.3. *Assume A1–A2.*

(i) *For all $\iota > 0$ such that $r - (1/2)\lambda_+\iota(1 - p) > 0$, $\int \pi(\mathrm{d}x)\exp(\iota|x|^{1-p}) < \infty$ where π is defined in Proposition 5.1.*
(ii) *There exists a closed petite set C such that for any $0 < m < 1 - p$, $0 < \iota_1 < \iota_2$ and $\delta > 0$, there exists a finite constant c such that*

$$\mathbb{E}_x\left[\exp(\iota_1\,\{\tau_C(\delta)\}^{\frac{m}{1+p}})\right] \le c\exp(\iota_2|x|^m). \tag{5.3}$$

If $m = 1 - p$, (5.3) still holds for any $0 < \iota_1 < \iota_2$ such that $r - (1/2)\iota_2\lambda_+(1 - p) > 0$.

This is a direct application of Proposition 3.1 and Theorem 4.1(ii) and the proof is omitted. The results of Theorem 5.3 can be compared to those by [19], where subexponential ergodicity in total variation norm of a diffusion satisfying the conditions A1–A2 is addressed, using a technique based on the coupling method. Theorem 5.3(i) states the same result as in [19, Lemma 3]. Nevertheless, Theorem 5.3(ii) yields a stronger control of delayed return-time to a closed petite set than those obtained in [19, Theorem 5]. They show that for all $0 < \alpha < (1/2)(1 - p)$ there exists a constant c_α such that $\mathbb{E}_x\left[\exp(\tau_C(\delta)^\alpha)\right] \le c_\alpha \exp(|x|^{2\alpha})$ and this remains valid for $\alpha = (1 - p)/2$ if $r - (1/2)\lambda_+(1 - p) > 0$. Theorem 5.3(ii) claims that for all $0 < \alpha < (1 - p)(1 + p)^{-1}$ and $\iota > 1$, $\mathbb{E}_x\left[\exp(\tau_C(\delta)^\alpha)\right] \le c_\alpha \exp(\iota\, |x|^{(1+p)\alpha})$ and for $\alpha = (1 - p)(1 + p)^{-1}$, $\mathbb{E}_x\left[\exp(\iota_1\tau_C(\delta)^\alpha)\right] \le c_\alpha\exp(\iota_2\, |x|^{(1+p)\alpha})$ for all $0 < \iota_1 < \iota_2$ such that $r - (1/2)\,\iota_2\lambda_+(1 - p) > 0$.

As a direct application of Theorem 3.2, we obtain the following results for f-ergodicity at a subgeometric rate.

Theorem 5.4. *Assume A1–A2 and let π be the invariant probability distribution of the Markov process that solves (5.1). Then the process is subgeometrically f-ergodic: for any $x \in \mathbb{R}^n$, the limits (3.4) to (3.7) hold with $V(x) \sim \exp(\iota|x|^{1-p})$ for some positive ι such that $r - 0.5\lambda_+\iota(1 - p) > 0$, $f_*(x) \sim |x|^{-2p}\exp(\iota|x|^{1-p})$ and $r_*(t) \sim t^{-2p/(1+p)}\exp(\{\iota' t\}^{(1-p)/(1+p)})$ where*

$$\iota' = \iota^{\frac{1+p}{1-p}}(1 + p)\,\{r - (1/2)\lambda_+\iota(1 - p)\}\,.$$

In [19], only the convergence in total variation norm of the semi-group $\{P^t\}_{t\ge 0}$ to the invariant probability π is addressed: it is established that the process is ergodic at the rate $r_*^{\mathrm{M}}(t) \propto \exp(\delta t^{(1-p)/2})$ for some $\delta > 0$, and in that case, the dependence upon the initial point in (3.4) is $V^{\mathrm{M}}(x) \sim \exp(\delta|x|^{1-p})$. Theorem 5.4 improves these results and also provides rates of convergence in f-norm for unbounded functions f.

We reported in Theorem 5.4 the values (V, f_*, r_*) that yield the best rate of convergence in total variation norm. Proposition 5.2 shows that one could establish the drift inequality (3.11) with $V(x) \sim \exp(\iota|x|^m)$ for some $0 < m < 1-p$; this would imply the limits (3.4) to (3.7) with $V(x) \sim \exp(\iota|x|^m)$, $f_*(x) \sim |x|^{m-1-p}\exp(\iota|x|^m)$ and $r_*(t) \sim t^{(m-1-p)/(1+p)}\exp(\iota'|t|^{m/(1+p)})$ for all $0 < \iota' < \iota$. We thus obtain a weaker maximal rate function r_*, and a weaker maximal norm $\|\cdot\|_{f_*}$, but this has to be balanced with the fact that the dependence upon the initial value (i.e. the quantity $V(x)$) is weaker too. Similarly, polynomially increasing controls $V(x)$ could be considered, thus limiting the rate r_* (resp. the function f_*) to the class of the polynomially increasing rate functions (resp. to the class of the polynomially increasing functions). These discussions illustrate the fact that the pair (ϕ, V) that solves (3.11) is not unique, and this results in balancing the pair (r_*, f_*) and the dependence upon the initial value x.

5.2. *Langevin tempered diffusions on* $\mathbb{R}^n$

Let $\pi : \mathbb{R}^n \to (0,\infty)$ satisfying

B1 π is, up to a normalizing constant, a positive and thrice continuously differentiable density on $\mathbb{R}^n$, with respect to the Lebesgue measure.

B2 there exists $0 < \beta < 1$ such that for all large $|x|$: $|x|^{1-\beta}\,\langle \partial \ln \pi(x), \frac{x}{|x|}\rangle < 0$ and

$$0 < \liminf_{x\to\infty} |\partial \ln \pi(x)||\ln \pi(x)|^{1/\beta-1} \le \limsup_{x\to\infty} |\partial \ln \pi(x)||\ln \pi(x)|^{1/\beta-1} < \infty,$$

$$\limsup_{x\to\infty} \mathrm{Tr}\left(\partial^2 \ln \pi(x)\right)|\partial \ln \pi(x)|^{-2} = 0.$$

Since π is defined up to a normalizing constant, we can assume that $\pi(x) < 1$ for all x ; set $\sigma(x) = |\ln \pi(x)|^d$ for some $d > 0$, define the diffusion matrix by $a(x) = \sigma^2(x)\mathrm{Id}_n$, and the drift vector by $b(x) = (b_1(x), \cdots, b_n(x))'$ where

$$b_i(x) = (1/2)\sum_{j=1}^n a_{ij}(x)\,\partial_{x_j} \ln \pi(x) + (1/2)\sum_{j=1}^n \partial_{x_j} a_{ij}(x), \quad 1 \le i \le n.$$

Our objective is to study the ergodicity of the solution to the stochastic integral equation

$$X_t = X_0 + \int_0^t b(X_s)\mathrm{d}s + \int_0^t \sigma(X_s)\mathrm{d}B_s \tag{5.4}$$

where $\{B_t\}_t$ is an n-dimensional Brownian motion. This diffusion is the so-called Langevin diffusion and the drift vector b is defined in such a way that π is, up to a multiplicative constant, the density of the unique invariant probability distribution. This model is not a particular case of the elliptic diffusion of Section 5.1 since here, σ may be an unbounded function ($\sigma = |\ln \pi(x)|^d$). Fort and Roberts investigate the behavior of these diffusions when π is polynomially decreasing in the tails and $\sigma(x) = \pi^{-d}(x)$ $(d > 0)$ [11]. We consider the case when π is subexponentially decreasing in the tails: the class of density π described by B1–B2 contains densities that are subexponential in the tails. The Weibull distribution on $(0,\infty)$ with density $\pi(x) \propto x^{\beta-1}\exp(-\alpha x^\beta)$ satisfies B2. For multi-dimensional examples, see e.g. [27,10].

The process is regular – whenever $d > 0$ – and there exists a solution to (5.4) which is an almost surely continuous stochastic process and is unique up to equivalence. This solution is an homogeneous strong Markov process whose transition functions are Feller functions. Furthermore, π is (up to a scaling factor) the density of an invariant distribution of the diffusion process, any skeleton chain is ψ-irreducible and compact sets are closed petite sets [11, Proposition 15]. Finally, for a twice continuously differentiable function $V : \mathbb{R}^n \to [1, \infty)$, $\mathcal{A}V(x) = LV(x)$ where L is the diffusion operator (5.2).

Set $V(x) = 1 + \pi^{-\kappa}(x)$ outside a compact set; standard computations yield, for large $|x|$,

$$\mathcal{A}V(x) \le -c_\kappa \, [\ln V(x)]^{-\alpha} \, V(x), \quad \text{where } \alpha = 2(1/\beta - 1 - d), \quad \text{and}$$
$$c_\kappa > 0 \Longleftrightarrow 0 < \kappa < 1.$$

If $\alpha \le 0$, the process is V-geometrically ergodic [20, Theorem 6.1] (see also Section 3.3); if $\alpha > 0$, it is subgeometrically ergodic as a consequence of Theorems 3.2 and 3.4. Geometric ergodicity was already observed [20, Theorem 6.1] while the subgeometric ergodicity is a new result.

Set $V(x) = 2 + \mathrm{sign}(\kappa)(-\ln \pi(x))^\kappa$ outside a compact set (we can assume w.l.g. that for large x, $\ln \pi(x) < 0$). Then there exists a constant $c > 0$ such that for large x,

$$\mathcal{A}V(x) \le -cV^{1-\alpha}(x), \quad \text{where } \alpha = 2\kappa^{-1}(1/\beta - d - (1/2)). \tag{5.5}$$

First consider the case when $\kappa > 0$. If $1/\beta - 1/2 - \kappa/2 < d < 1/\beta - (1/2)$, the drift condition (5.5) and Theorems 3.2 and 3.4 yield polynomial ergodicity. For example, this implies convergence in total variation norm at the rate $r(t) \sim t^{1/\alpha - 1}$. If $d = 1/\beta - (1/2)$, then $\alpha = 0$ and the process is geometrically ergodic. In the case when κ can be set negative and $1 - \alpha > 0$ i.e. when $d > 1/\beta - (1/2)$, the process is uniformly ergodic: there exist $\lambda < 1$ and a constant $c < \infty$ such that for all x, $\lambda^{-t} \, \|P^t(x, \cdot) - \pi(\cdot)\|_{\mathrm{TV}} \le c$, and the convergence does not depend on the starting point.

The above discussions are summarized in the following theorem. The first part (resp. third part) results from [29, Theorem 3.1] (resp. [20, Theorem 6.1]). The second assertion is a consequence of Theorem 3.2. The last assertion was already proved by [29, Theorem 3.1] for one-dimensional diffusions ($n = 1$).

Theorem 5.5. *Consider the Langevin diffusion on $\mathbb{R}^n$ solution to the Eq. (5.4) where the target distribution π satisfies B1–B2.*

(i) *If $0 \le d < 1/\beta - 1$, the process fails to be geometrically ergodic.*
(ii) *If $0 \le d < 1/\beta - 1$, the process is subgeometrically ergodic: the limits (3.4) to (3.7) hold with $V(x) \sim \pi^{-\kappa}(x)$, $f_*(x) \sim \pi^{-\kappa}(x) \, |\ln \pi(x)|^{-2(1/\beta - 1 - d)}$ and $\ln r_*(t) \sim c_\kappa t^{\beta/(2-\beta-2d\beta)}$ for all $0 < \kappa < 1$.*
(iii) *If $d \ge 1/\beta - 1$, then for all $0 < \kappa < 1$, the diffusion is V-geometrically ergodic with $V(x) = 1 + \pi^{-\kappa}(x)$.*
(iv) *If $d > 1/\beta - (1/2)$, the diffusion is uniformly ergodic.*

This theorem extends earlier results for the multi-dimensional case and provides subgeometrical rates of convergence of the 'cold' Langevin diffusion, for a wide family of norms. We established that for a given $\pi^{-\kappa}$-norm, the minimal rate of convergence is achieved with $d = 0$ and in that case, the rate coincides with the rate of convergence of the symmetric random-walk Hastings–Metropolis algorithm ([5, Theorem 3.1]). This rate can be improved by choosing a diffusion matrix which is heavy where π is light and conversely. When d is larger than the critical value $d_* = 1/\beta - 1$, the process is geometrically ergodic; when d is lower than d_*, the process cannot be geometrically ergodic and we prove that it is subgeometrically ergodic. The conclusions of Theorem 5.5 are similar to those of [11, Theorem 16] who address the case when π is polynomial in the tails.

We assumed that $\sigma = |\ln \pi|^d$, $d > 0$. A first extension is to consider a sufficiently smooth function σ such that $\sigma(x) \sim |\ln \pi(x)|^d$ for large $|x|$; this yields similar conclusions and the details are omitted. A second extension consists of the case when $\sigma(x) \sim \pi^{-d}(x)$. In this latter case, following the same lines, it is easily verified that for small enough values of d, the process is regular (the set of the admissible values is in the range $(0, 1/2]$), and the process is V-geometrically ergodic with a test function $V(x) \sim \pi^{-\kappa}(x)$, $\kappa > 0$. The details are omitted and left to the interested reader.

5.3. Stochastic damping Hamiltonian system

Both examples in the previous sections assumed that the diffusion process is elliptic. However the drift condition (3.11) enables us to consider also hypoelliptic diffusion that we will illustrate on the example of a simple stochastic damping Hamiltonian system, i.e. let x_t (resp. y_t) be the position (resp. the velocity) at time t of a physical system moving in $\mathbb{R}^n$

$$\begin{aligned} \mathrm{d}X_t &= Y_t \mathrm{d}t \\ \mathrm{d}Y_t &= \Sigma(X_t, Y_t)\mathrm{d}B_t - (c(X_t, Y_t)Y_t + \partial_x U(X_t))\mathrm{d}t \end{aligned} \tag{5.6}$$

where $-\partial_x U$ is some friction force, $-c(x, y)y$ is the damping force and $\Sigma(x, y)\mathrm{d}B$ is a random force where (B_t) is a standard Brownian motion in $\mathbb{R}^n$. This system has been studied from the large and moderate deviation points of view by Wu [32] where he also establishes the exponential ergodicity under various sets of assumptions.

As our goal is not to consider the model in its full generality but to illustrate the subexponential behavior of hypoelliptic diffusion, via the simple use of drift condition (3.11), hereafter we will consider the particular one-dimensional case

$$\mathrm{d}X_t = Y_t \mathrm{d}t \quad \mathrm{d}Y_t = \sigma \mathrm{d}B_t - (cY_t + U'(X_t))\mathrm{d}t, \tag{5.7}$$

and assume that U is C^2, and there exist $0 < p < 1$ and positive constants a, b such that for $|x|$ large enough

$$a|x|^{p-1} \leq U'(x) \leq b|x|^{p-1}. \tag{5.8}$$

By Wu [32, Lemma 1.1, Proposition 1.2], the solution is a strong Markov process, all the skeletons are irreducible and compact sets are petite sets and admits an unique invariant probability measure

$$\pi(\mathrm{d}x, \mathrm{d}y) = \mathrm{e}^{-\frac{2c}{\sigma}H(x,y)}\mathrm{d}x\mathrm{d}y$$

where H is the Hamiltonian given by $H(x, y) = \frac{1}{2}|y|^2 + U(x)$.

The fact that p is less than 1 implies that $((X_t, Y_t))_{t\geq 0}$ cannot be exponentially ergodic [32, Theorem 5.1]. We now exhibit a drift function satisfying (3.11). Consider positive constants α, β and a smooth positive function G such that for m, $1 - p < m \leq 1$, $G'(x) = |x|^m$ for large $|x|$; define a twice continuously differentiable function $V_m \geq 1$ such that for large x, y,

$$V_m(x, y) = \alpha(y^2/2 + U(x)) + \beta(G'(x)y + cG(x)).$$

By the definition of $\mathcal{A}$, it holds

$$\mathcal{A}V_m(x,y) = \frac{1}{2}\sigma^2\, \partial_y^2 V_m(x,y) + y\, \partial_x V_m(x,y) - (cy + U'(x))\partial_y V_m(x,y)$$

so that

$$\begin{aligned}\mathcal{A}V_m(x,y) &= \frac{1}{2}\alpha\sigma^2 + y(\alpha U'(x) + \beta G''(x)y + \beta c G'(x)) - (cy + U'(x))(\alpha y + \beta G'(x)) \\ &= \frac{1}{2}\alpha\sigma^2 + (\beta G''(x) - c\alpha)y^2 - \beta G'(x)U'(x).\end{aligned}$$

Fix $\delta < 0$; since $m \leq 1$, we choose β small enough so that $\beta G''(x) - c\alpha < \delta < 0$ for all large x. Furthermore, for all large $|x|$, $G'(x)U'(x) \geq a|x|^{p-1+m}$. Hence, there exist positive constants K, L such that

$$\mathcal{A}V_m(x,y) \leq K - L\ V_m(x,y)^{\frac{p-1+m}{m+1}}.$$

Condition (3.11) holds with $\phi_m(v) \propto v^{\frac{p-1+m}{m+1}}$ and $\frac{p-1+m}{m+1} < 1$. The application of the results of Section 3.1 now implies that the process $(Z_t)_{t\geq 0}$ is polynomially ergodic.

Let $k \geq 1$ and define a twice continuously differentiable function $V_{m,k} \geq 1$ such that for $|x|^2 + |y|^2$ sufficiently large

$$V_{m,k}(x,y) = V_m^k(x,y).$$

Then for $|x|^2 + |y|^2$ sufficiently large, the above calculations yield

$$\begin{aligned}\mathcal{A}V_{m,k}(x,y) &= k(\mathcal{A}V_m(x,y))V_m^{k-1}(x,y) + \frac{k(k-1)}{2}\sigma^2(\partial_y V_m(x,y))^2 V_m^{k-2}(x,y) \\ &= \left(k(\mathcal{A}V_m(x,y)) + \frac{k(k-1)}{2}\sigma^2\frac{(\partial_y V_m(x,y))^2}{V_m(x,y)}\right)V_m^{k-1}(x,y) \\ &\leq (K' - LV_m(x,y)^{\frac{p-1+m}{m+1}})V_m^{k-1}(x,y) \\ &\leq K'' - L'V_m^{\frac{p-1+m}{m+1}+k-1}\end{aligned}$$

for some positive constant K', K'', L'. This inequality is once again the condition (3.11) with $\phi_{m,k}(v) = v^{(\frac{p-2}{m+1}+k)k^{-1}}$. These discussions are summarized in the following theorem.

Theorem 5.6. *Let U be a twice continuously differentiable function, lower bounded on $\mathbb{R}$ satisfying (5.8) for some $0 < p < 1$. Then $(Z_t)_{t\geq 0}$ is not exponentially ergodic but is polynomially ergodic: for any m such that $1 - p < m \leq 1$ and any $k \geq 1$, the limits (3.4)–(3.7) hold with V_{mk} defined above, $\phi_{m,k}(v) \propto v^{(\frac{p-2}{m+1}+k)k^{-1}}$, $f_* = \phi_{m,k} \circ V_{m,k}$ and $r_*(t) \propto t^{\frac{k(m+1)}{2-p}-1}$.*

Observe that the process $((X_t, Y_t))_{t\geq 0}$ is polynomially ergodic at any order and we strongly believe it is subexponentially ergodic. This example shows that our conditions are sufficiently flexible to consider the hypoelliptic diffusions as well as the elliptic ones.

5.4. Compound Poisson-process driven Ornstein–Uhlenbeck process

In this section we consider an example of Fort–Roberts [11] where subgeometric ergodicity can be achieved where they only obtain polynomial ergodicity. Let us first recall the model. Let

X be an Ornstein–Uhlenbeck process driven by a finite rate subordinator:

$$\mathrm{d}X_t = -\mu X_t + \mathrm{d}Z_t$$

and $Z_t = \sum_{i=1}^{N_t} U_i$, where $(U_i)_{i\geq 1}$ is a sequence of non-negative i.i.d. r.v. with probability measure F, and (N_t) is an independent Poisson process of rate λ. We suppose the recall coefficient μ to be positive. Remarking that only when F is sufficiently (even extremely) heavy tailed, X fails to be exponentially ergodic, Fort–Roberts [11] give conditions for which X is polynomially ergodic. Namely, denote by G the law of the log jump sizes ($G(A) = F(\mathrm{e}^A)$), and assume that for all $\kappa > 0$, $\int \mathrm{e}^{\kappa x}\mathrm{d}G(x) = +\infty$. Lemma 17 of Fort–Roberts then prove that X is not exponentially ergodic and give examples where X is positive recurrent and polynomially ergodic, namely when for some $r > 1$, $\int_0^\infty [\ln(1+u)]^r F(\mathrm{d}u)$ is finite. Such an assertion may be useful considering for large x (C the normalizing constant)

$$F(\mathrm{d}x) = \frac{C_k^{-1}}{x(\ln(x))^k}\mathrm{d}x \quad k > 1 \qquad F(\mathrm{d}x) = \frac{C_{\beta,c}^{-1}\mathrm{e}^{-c(\ln(x))^\beta}}{x}\mathrm{d}x \quad 0 < \beta \leq 1.$$

We shall strengthen their result by:

Proposition 5.7. *Suppose that (X_t) is aperiodic and that for some $\delta < 1$, $\alpha > 0$*

$$\int_0^\infty \mathrm{e}^{\alpha(\ln(1+x))^\delta} F(\mathrm{d}x) < \infty.$$

Then, the conclusions of Theorem 3.2 hold with $V(x) = \mathrm{e}^{\alpha'(\ln x)^{\delta'}}$ for $\delta' < \delta$ and positive α' (and $\alpha' < \alpha$ if $\delta' = \delta$), and $\phi(v) = cv/\log(v)^{(1-\delta')/\delta'}$ with $c = \alpha'^{2-\delta'}\mu\delta'$, $r_*(t) = t^{-(1+\delta')}\mathrm{e}^{(ct/\delta')^{\delta'}}$, $f_* = \phi \circ V$.

Proof. We shall use the drift conditions introduced previously for the generator defined for all functions V in the extended domain of the generator

$$\mathcal{A}V(x) = \lambda \int_0^\infty (V(x+u) - V(x))F(\mathrm{d}u) - \mu x V'(x).$$

Choosing $V(x) = (\ln(x))^r$, as in Fort–Roberts [11, Lemma 18], for sufficiently large x ensures the polynomial ergodicity. Consider now $V(x) = \mathrm{e}^{\alpha'(\ln x)^{\delta'}}$ for large x, so that

$$\begin{aligned}\mathcal{A}V(x) &= \lambda \int_0^\infty (\mathrm{e}^{\alpha'(\ln(x+u))^{\delta'}} - \mathrm{e}^{\alpha'(\ln x)^{\delta'}})F(\mathrm{d}u) - \alpha'\delta'\mu \frac{\mathrm{e}^{\alpha'(\ln x)^{\delta'}}}{(\ln x)^{1-\delta'}} \\ &\leq -\alpha'^{2-\delta'}\mu\delta' \frac{V}{(\ln V)^{(1-\delta')/\delta'}} + b\end{aligned}$$

using for the integral term, dominated convergence term, and that for large x, $\mathrm{e}^{\alpha'(\ln(x+u))^{\delta'}} - \mathrm{e}^{\alpha'(\ln x)^{\delta'}} \sim \delta' \frac{\mathrm{e}^{\alpha'(\ln x)^{\delta'}}}{(\ln x)^{1-\delta'}} \ln(1+u/x)$. $\square$

Acknowledgments

The authors sincerely thank the referees and the editor for many comments enabling them to set a better presentation. This work was partly supported by the National Research Agency (ANR) under the program "ANR-05-BLAN-0299".

Appendix. Proofs

Proof of Theorem 4.1. (i) is a direct application of the optional sampling theorem for the right-continuous supermartingale (3.2) (see e.g. [8, Theorem 2.13 p. 61]) with the bounded $\mathcal{F}$-stopping time $\tau = \tau_C(\delta) \wedge M$ and by letting $M \to \infty$.

(ii) Let $G(t,u) = H_\phi^{-1}(H_\phi(u)+t) - H_\phi^{-1}(t)$. Note that

$$\begin{aligned}\frac{\partial G(t,u)}{\partial u} &= \frac{\phi \circ H_\phi^{-1}(H_\phi(u)+t)}{\phi(u)},\\ \frac{\partial G(t,u)}{\partial t} &= \phi \circ H_\phi^{-1}(H_\phi(u)+t) - \phi \circ H_\phi^{-1}(t).\end{aligned} \tag{A.1}$$

By log-concavity of $\phi \circ H_\phi^{-1}$, for any fixed t, $u \mapsto \frac{\partial G(t,u)}{\partial u}$ is non-increasing and thus, for any fixed t, the function $u \mapsto G(t,u)$ is concave.

Set $\tau = \tau_C(\delta) \wedge \tau_m$ with $\tau_m = \inf\{s \geq 0, |X_s| \geq m\}$. Let $\epsilon > 0$. Write $t_k = \epsilon k$ and set $N_\epsilon = \sup\{k \geq 1; t_{k-1} < \tau\}$. Note that by (i), $\mathbb{P}_x(\tau_C(\delta) < \infty) = 1$. Furthermore, $\tau_C(\delta) \leq \epsilon N_\epsilon$ and ϵN_ϵ is an $\mathcal{F}$-stopping time. This implies that for any $M > \delta$,

$$\begin{aligned}\mathbb{E}_x\left[\int_0^{\tau\wedge M} \phi \circ H_\phi^{-1}(s)\mathrm{d}s\right] &\leq \limsup_{\epsilon\to 0} \mathbb{E}_x\left[\int_0^{(\epsilon N_\epsilon)\wedge M} \phi \circ H_\phi^{-1}(s)\mathrm{d}s\right]\\ &= \limsup_{\epsilon\to 0} \mathbb{E}_x\left[\int_0^{\epsilon(N_\epsilon\wedge M_\epsilon)} \phi \circ H_\phi^{-1}(s)\mathrm{d}s\right] \leq \limsup_{\epsilon\to 0} A(\epsilon) + G(0, V(x))\end{aligned} \tag{A.2}$$

where we set $M_\epsilon = \lfloor M/\epsilon \rfloor$ ($\lfloor\cdot\rfloor$ denotes the lower integer part) and

$$\begin{aligned}A(\epsilon) = {}& \mathbb{E}_x\left[G(\epsilon(N_\epsilon \wedge M_\epsilon), V(X_{\epsilon(N_\epsilon\wedge M_\epsilon)})) - G(0, V(x))\right]\\ &+ \mathbb{E}_x\left[\int_0^{\epsilon(N_\epsilon\wedge M_\epsilon)} \phi \circ H_\phi^{-1}(s)\mathrm{d}s\right].\end{aligned}$$

The proof is completed provided it is established that $\limsup_{\epsilon\to 0} A(\epsilon) \leq \frac{b}{\phi(1)} \int_0^\delta \phi \circ H_\phi^{-1}(s)\mathrm{d}s$ which we now prove.

$$\begin{aligned}&A(\epsilon) - \mathbb{E}_x\left[\int_0^{\epsilon(N_\epsilon\wedge M_\epsilon)} \phi \circ H_\phi^{-1}(s)\mathrm{d}s\right]\\ &\quad= \mathbb{E}_x\left[\sum_{k=1}^{M_\epsilon} \left\{G(t_k, V(X_{t_k})) - G(t_{k-1}, V(X_{t_{k-1}}))\right\} \mathbf{1}_{\tau > t_{k-1}}\right]\\ &\quad\leq \mathbb{E}_x\left[\sum_{k=1}^{M_\epsilon} \mathbb{E}\left[G(t_k, V(X_{t_k})) - G(t_{k-1}, V(X_{t_{k-1}}))\middle|\mathcal{F}_{t_{k-1}}\right] \mathbf{1}_{\tau > t_{k-1}}\right]\end{aligned} \tag{A.3}$$

where we have used that $\{\tau > t_{k-1}\} \in \mathcal{F}_{t_{k-1}}$. Moreover, by concavity of $u \to G(t,u)$,

$$\begin{aligned}&\mathbb{E}\left[G(t_k, V(X_{t_k})) - G(t_{k-1}, V(X_{t_{k-1}}))\middle|\mathcal{F}_{t_{k-1}}\right]\\ &\quad\leq \frac{\partial G}{\partial u}(t_k, V(X_{t_{k-1}}))\mathbb{E}\left[V(X_{t_k}) - V(X_{t_{k-1}})\middle|\mathcal{F}_{t_{k-1}}\right] + \int_{t_{k-1}}^{t_k} \frac{\partial G}{\partial t}(s, V(X_{t_{k-1}}))\mathrm{d}s.\end{aligned}$$

(A.1) and (A.3), $\mathbf{D}(\mathbf{C}, \mathbf{V}, \phi, \mathbf{b})$ and the log-concavity of $\phi \circ H_\phi^{-1}$ yield

$$A(\epsilon) \leq \mathbb{E}_x \left[\sum_{k=1}^{M_\epsilon} \phi \circ H_\phi^{-1}(H_\phi(V(X_{t_{k-1}})) + t_k) \left(-\frac{\int_{t_{k-1}}^{t_k} \phi \circ V(X_s) \mathrm{d}s}{\phi(V(X_{t_{k-1}}))} + \epsilon \right) \mathbf{1}_{\tau > t_{k-1}} \right] + \frac{b}{\phi(1)} \mathbb{E}_x \left[\int_0^{\epsilon(N_\epsilon \wedge M_\epsilon)} \phi \circ H_\phi^{-1}(s+\epsilon) \mathbf{1}_C(X_s) \mathrm{d}s \right].$$

Consider the first term of the rhs. Define

$$g = \phi \circ H_\phi^{-1}(H_\phi(\sup_{t \in [0,M]} V(X_t)) + M) \left| \frac{\int_0^M \phi \circ V(X_s) \mathrm{d}s}{\phi(1)} + M \right|. \tag{A.4}$$

Applying a Fatou-lemma-type inequality of the form $\limsup_n \mathbb{E}[f_n] \leq \mathbb{E}[\limsup_n f_n]$ for functions $\{f_n\}_n$ satisfying $|f_n| \leq g$ and $\mathbb{E}[g] < \infty$ where g is defined in (A.4)

$$\begin{aligned}
&\limsup_{\epsilon \to 0} \mathbb{E}_x \left[\sum_{k=1}^{M_\epsilon} \phi \circ H_\phi^{-1}(H_\phi(V(X_{t_{k-1}})) + t_k) \left| -\frac{\int_{t_{k-1}}^{t_k} \phi \circ V(X_s) \mathrm{d}s}{\phi(V(X_{t_{k-1}}))} + \epsilon \right| \right] \\
&\quad \leq \mathbb{E}_x \left[\phi \circ H_\phi^{-1}(H_\phi(\sup_{t \in [0,M]} V(X_t)) + M) \right. \\
&\quad \times \left. \limsup_{\epsilon \to 0} \sum_{k=1}^{M_\epsilon} \left| \frac{\int_{t_{k-1}}^{t_k} \{\phi \circ V(X_s) - \phi \circ V(X_{t_{k-1}})\} \mathrm{d}s}{\phi(1)} \right| \right].
\end{aligned}$$

This term is zero since for any cad-lag function g, $\limsup_{\epsilon \to 0} \sum_{k=1}^{M_\epsilon} \int_{t_{k-1}}^{t_k} |g(s) - g(t_{k-1})| \mathrm{d}s = 0$ [1, Chapter 3]. Thus, using again the Fatou-lemma-type inequality,

$$\begin{aligned}
\frac{\phi(1)}{b} \limsup_{\epsilon \to 0} A(\epsilon) &\leq \limsup_{\epsilon \to 0} \mathbb{E}_x \left[\int_0^{\epsilon(N_\epsilon \wedge M_\epsilon)} \phi \circ H_\phi^{-1}(s+\epsilon) \mathbf{1}_C(X_s) \mathrm{d}s \right] \\
&\leq \mathbb{E}_x \left[\int_0^M \phi \circ H_\phi^{-1}(s) \mathbf{1}_C(X_s) \left(\limsup_{\epsilon \to 0} \mathbf{1}_{s \leq \epsilon N_\epsilon < \tau + \epsilon} \right) \mathrm{d}s \right] \\
&\leq \mathbb{E}_x \left[\int_0^M \phi \circ H_\phi^{-1}(s) \mathbf{1}_C(X_s) \mathbf{1}_{s \leq \tau} \mathrm{d}s \right] \leq \int_0^\delta \phi \circ H_\phi^{-1}(s) \mathrm{d}s.
\end{aligned}$$

The proof follows by letting $m, M \to \infty$. $\square$

Proof of Proposition 4.4. The ψ-irreducibility results from [20, Theorem 1.1].

Let A be a closed accessible petite set, the existence of which is proved in [20, Proposition 3.2(i)]. B is petite provided $\sup_{x \in B} \mathbb{E}_x[\tau_A] < +\infty$ (see [20, Proposition 4.2]). As shown in the proof of [20, Proposition 4.1], there exist $\delta > 0$ and $c_1 < \infty$ such that for all $x \in \mathsf{X}$, $\mathbb{E}_x[\tau_A] \leq \mathbb{E}_x[\tau_C(\delta)] + c_1$. The proof is completed by applying Theorem 4.1.

We can assume w.l.g. that C is ν_a-petite and ν_a is a maximal irreducibility measure [20, Proposition 3.2(ii)]. Since $\cup_n B_n$ is full, there exists n_* such that $C \subseteq B_{n_*}$ and B_{n_*} is accessible. Since B_{n_*} is accessible and ν_a is regular, there exists a compact set $B \subseteq B_{n_*}$ such that $\nu_a(B) > 0$. Then B is closed, accessible, petite and $\sup_B V < n_*$. $\square$

Proof of Proposition 4.5. (i) The proof is along the same lines as the proof of [20, Proposition 4.2] upon noting that (a) by [11, Lemma 20], there exists $M < \infty$ such that for all $t \geq 0$,

$G_C(\cdot, f, r; \delta + t) \leq G_C(\cdot, f, r; \delta) + M^t$; and (b) we can assume that C is ν_a-petite for some maximal irreducibility measure ν_a and a distribution a such that $\int M^t a(\mathrm{d}t) < \infty$ [20, Proposition 3.2(ii)].

(ii) We can assume without loss of generality that $r \in \Lambda_0$ and we will do so. Assume that for any $t \geq t_0$, there exists a constant $c_t < +\infty$ such that $G_B(\cdot, f, r; t) \leq c_t\ G_C(\cdot, f, r; t)$. We then apply [11, Lemma 20] and the proof is concluded. We now consider the construction of such a constant c_t. The proof is along the same lines as [25, Lemma 3.1] and we only give the sketch of the proof: there exist $\gamma > 0$ and $t_0 > 0$ such that $\inf_{x \in C} \mathbb{P}_x(\tau_B \leq t_0) \geq \gamma$ since C is petite and B is accessible. Let $t \geq t_0$. Set $\tau = \tau_C(t)$ and denote by τ^k the kth-iterate of τ: $\tau^{k+1} = \tau^k + \tau \circ \theta^{\tau^k}$ for any $k \geq 1$, where θ is the usual shift operator. Define for $n \geq 2$, the $\{0, 1\}$-valued random variables $(u_n)_n$ by $u_n = 1$ iff $\tau_B \circ \theta^{\tau^{n-1}} \leq t$. Then by definition, $u_n \in \mathcal{F}_{\tau^n}$ and $\mathbb{P}_x\left(u_n = 1 | \mathcal{F}_{\tau^{n-1}}\right) \geq \gamma$. Finally, set $\eta = \inf\{n \geq 2, u_n = 1\}$. Then it holds

$$G_B(x, f, r; t) \leq \mathbb{E}_x\left[\int_0^{\tau^{\eta-1}+t} r(s) f(X_s)\, \mathrm{d}s\right] \leq \sum_{n \geq 2} (a_x(n) + M_t\, b_x(n))$$

where we set $M_t = \sup_C G_C(\cdot, f, r; t)$, $a_x(n) = \mathbb{E}_x\left[\int_0^{\tau^{n-1}} r(s) f(X_s)\, \mathrm{d}s\, \mathbf{1}_{\eta \geq n}\right]$ and $b_x(n) = \mathbb{E}_x\left[r(\tau^{n-1}) \mathbf{1}_{\eta \geq n}\right]$. Following the same lines as in the proof of [25, Lemma 3.1], it may be proved that for all $n \geq 2$, $b_x(n) \leq \rho\ b_x(n-1) + c\ (1-\gamma)^{n-1}$ and $a_x(n) \leq (1-\gamma)\ a_x(n-1) + M_t\ b_x(n-1)$ for some constants $0 < c < \infty$ and $0 < \rho < 1$. This proves that there exists a constant $c_t < \infty$ such that $G_B(\cdot, f, r; t) \leq c_t\ G_C(\cdot, f, r; t)$. □

Proof of Proposition 4.7. (ii) ⇒ (i) is trivial. (i) ⇒ (ii): Proposition 4.5(ii) implies that for all $n \geq 1$, the set $\{x \in \mathsf{X}, G_C(x, f, r; \delta) \leq n\}$ is (f, r)-regular and thus petite ([20, Proposition 4.2(i)]); furthermore, their union is full (Proposition 4.5(i)). Since ψ is regular, we then conclude as in the proof of Proposition 4.4 that there exists a (f, r)-regular set, which is petite, closed and accessible.

(i) ⇒ (iii): set $\mathcal{S}_\Psi = \{x \in \mathsf{X}, G_C(x, f, r; \delta) < +\infty\}$.

(iii) ⇒ (ii): the proof is similar to the proof of Proposition 4.4 since the measure ψ is regular. □

Proof of Proposition 4.8. (i) We first establish that there exists a finite constant c such that $\mathbb{E}_x[T_{m,B}] \leq c\mathbb{E}_x[\tau_C(\delta)]$. To attain that goal, observe that the process is positive Harris-recurrent [20, Theorem 1.2] and some skeleton is irreducible so that, by [22, Proposition 6.1] and [20, Proposition 3.2(ii)], there exists $t_0 > 0$ s.t. $\inf_{x \in C} \inf_{t \in [t_0, t_0+m]} P^t(x, B) > 0$. The constant c can now be defined along the same lines as in the proof of [11, Proposition 22(ii)].

For any positive integer $M > 0$,

$$\begin{aligned}
&\mathbb{E}_x\left[\sum_{k=1}^{T_{m,B} \wedge M} \phi \circ V(X_{mk})\right] - \mathbb{E}_x\left[\int_0^{T_{m,B} \wedge M} \phi \circ V(X_{ms}) \mathrm{d}s\right] \\
&= \mathbb{E}_x\left[\sum_{k=1}^{\infty}\left[\int_{k-1}^{k} \{\phi \circ V(X_{mk}) - \phi \circ V(X_{ms})\} \mathrm{d}s\right] \mathbf{1}_{k \leq T_{m,B} \wedge M}\right] \\
&\leq \mathbb{E}_x\left[\sum_{k=1}^{\infty}\left[\int_{k-1}^{k} \{\phi' \circ V(X_{ms})\ (V(X_{mk}) - V(X_{ms}))\} \mathrm{d}s\right] \mathbf{1}_{k \leq T_{m,B} \wedge M}\right]
\end{aligned}$$

$$\leq \mathbb{E}_x\left[\sum_{k=1}^{\infty}\int_{k-1}^{k} \mathbb{E}_x\left[\,V(X_{mk}) - V(X_{ms})|\,\mathcal{F}_{ms}\right)\ \phi' \circ V(X_{ms}]\ \mathbf{1}_{k\leq T_{m,B}\wedge M}\right]\mathrm{d}s$$

$$\leq b\phi'(1)\mathbb{E}_x\left[\sum_{k=1}^{\infty}\int_{k-1}^{k}\int_{sm}^{km} \mathbf{1}_C(X_u)\mathrm{d}u\ \mathrm{d}s\ \mathbf{1}_{k\leq T_{m,B}\wedge M}\right] \leq b\phi'(1)\mathbb{E}_x\left[m(T_{m,B}\wedge M)\right].$$

This yields

$$\mathbb{E}_x\left[\int_0^{T_{m,B}} \phi\circ V(X_{ms})\mathrm{d}s\right] \leq \left(1 + b\frac{\phi'(1)}{\phi(1)}\right)\ \mathbb{E}_x\left[\int_0^{T_{m,B}} \phi\circ V(X_{ms})\mathrm{d}s\right].$$

The drift condition $\mathbf{D}(\mathbf{C}, \mathbf{V}, \phi, \mathbf{b})$ and the optional sampling theorem imply

$$\begin{aligned} m\mathbb{E}_x\left[\int_0^{T_{m,B}\wedge M} \phi\circ V(X_{ms})\mathrm{d}s\right] &= \mathbb{E}_x\left[\int_0^{m(T_{m,B}\wedge M)} \phi\circ V(X_s)\mathrm{d}s\right] \\ &\leq\ V(x) + bm\ \mathbb{E}_x\left[T_{m,B}\right]. \end{aligned} \tag{A.5}$$

The inequality $\mathbb{E}_x[T_{m,B}] \leq c\mathbb{E}_x[\tau_C(\delta)]$ and Theorem 4.1 yield the desired result.

(ii) Since $r_* = \phi\circ H_\phi^{-1}$ is increasing,

$$\mathbb{E}_x\left[\sum_{k=0}^{T_{m,B}-1} r_*(km)\right] \leq \phi(1) + \mathbb{E}_x\left[\int_0^{mT_{m,B}} r_*(s)\mathrm{d}s\right].$$

The result now follows from Theorem 4.1 and [11, Proposition 22(ii)], with a minor modification: the authors claim that $T_{m,B} \leq \tau^\eta$ while we have $mT_{m,B} \leq \tau_C^\eta$. □

Proof of Theorem 4.9. (i). A petite set C for the resolvent kernel, is also petite for the Markov process with semi-group P_t. By the definition of the resolvent kernel,

$$\mathbb{E}_x[R_\beta V(X_u)] = \int_0^\infty \beta\mathrm{e}^{-\beta v}P^{v+u}V(x)\ \mathrm{d}v = \mathrm{e}^{\beta u}R_\beta V(x) - \mathrm{e}^{\beta u}\int_0^u \beta\mathrm{e}^{-\beta v}P^v V(x)\ \mathrm{d}v\ .$$

This implies that

$$\begin{aligned} &\beta\ \mathbb{E}_x\left[\int_0^s \{R_\beta V(X_u) - V(X_u)\}\ \mathrm{d}u\right] \\ &= \int_0^s \beta\mathrm{e}^{\beta u}R_\beta V(x)\mathrm{d}u - \int_0^s\left(\mathrm{e}^{\beta u}\int_0^u \beta^2\mathrm{e}^{-\beta v}P^v V(x)\ \mathrm{d}v\right)\mathrm{d}u - \beta\int_0^s P^u V(x)\ \mathrm{d}u \\ &= (\mathrm{e}^{\beta s}-1)R_\beta V(x) - \beta\int_0^s\left(\int_v^s \beta\mathrm{e}^{\beta u}\mathrm{d}u\right)\mathrm{e}^{-\beta v}P^v V(x)\ \mathrm{d}v - \beta\int_0^s P^u V(x)\ \mathrm{d}u \\ &= (\mathrm{e}^{\beta s}-1)R_\beta V(x) - \mathrm{e}^{\beta s}\beta\int_0^s \mathrm{e}^{-\beta v}P^v V(x)\ \mathrm{d}v = \mathbb{E}_x[R_\beta V(X_s)] - R_\beta V(x). \end{aligned}$$

The proof follows observing that under $\check{\mathbf{D}}(\mathbf{C}, \mathbf{V}, \phi, \mathbf{b}, \beta)$

$$\mathbb{E}_x\left[\int_0^s \{R_\beta V(X_u) - V(X_u)\}\ \mathrm{d}u\right] \leq -\mathbb{E}_x\left[\int_0^s \phi\circ V(X_u)\mathrm{d}u\right] + b\mathbb{E}_x\left[\int_0^s \mathbf{1}_C(X_u)\mathrm{d}u\right].$$

(ii) By [20, Theorem 2.3(i) and Proposition 4.4(ii)] and Theorem 4.1, there exist positive constants δ, c_1 and c_2 such that for any $x \in \mathsf{X}$,

$$\check{\mathbb{E}}_x \left[\sum_{k=1}^{\check{\tau}_C} \phi \circ V(\check{X}_k) \right] \leq G_C(x, \phi \circ V, \mathbf{1}; \delta) + c_1 \sup_{x \in C} G_C(x, \phi \circ V, \mathbf{1}; \delta) \leq V(x) + c_2 \tag{A.6}$$

where $(\check{X}_k)_k$ is a Markov chain with transition kernel R_β, $\check{\tau}_C = \inf\{k \geq 1 : \check{X}_k \in C\}$ and $\check{\mathbb{E}}_x$ is the expectation associated to $\check{\mathbb{P}}_x$ the probability induced by the Markov chain $(\check{X}_k)_k$. Observe that $c_2 \geq \sup_C \phi \circ V$.

Define $W(x) = \check{\mathbb{E}}_x \left(\sum_{k=0}^{\check{\sigma}_C} \phi \circ V(\check{X}_k) \right)$ where $\check{\sigma}_C = \inf\{k \geq 0 : \check{X}_k \in C\}$; by (A.6) and the concavity of ϕ

$$W \leq \phi \circ V + V + c_2 \leq \left(\phi(1) + c_2 - \phi'(1) \right) + V(x) \left(1 + \phi'(1) \right). \tag{A.7}$$

Finally set $\check{C} = \{x \in \mathsf{X} : W(x) \leq c_2 + 1 + \phi(1)\}$ and let $\check{\phi}$ be a non-decreasing differentiable concave function such that $\check{\phi}(u) = \phi \left(\frac{u - [\phi(1) + c_2 - \phi'(1)]}{1 + \phi'(1)} \right)$ for $u \geq c_2 + 1 + \phi(1)$. From the equality $R_\beta W = W - \phi \circ V$ and the upper bound (A.7), we have

$$R_\beta W - W \leq -\check{\phi}(W) + \mathbf{1}_{\check{C}} \left(\check{\phi}(W) - \phi \circ V \right) \leq -\check{\phi}(W) + \sup_C \left(\check{\phi}(W) + \phi \circ V \right) \mathbf{1}_{\check{C}}.$$

Finally $\check{C} \subset \{x \in \mathsf{X} : \phi \circ V \leq c_2 + 1 + \phi(1)\}$ which is a level set of V since ϕ is increasing and differentiable. By Proposition 4.4, $\check{C}$ is petite for the process, and thus also, for the kernel R_β. □

Lemma A.1. *Let $r \in \Lambda_0$ and for any integer $m > 0$, define the sequence on the positive integers $\Delta^{(m)} r$ by $\Delta^{(m)} r(0) = 0$ and $\Delta^{(m)} r(k) = r(mk) - r(m(k-1))$, $k \geq 1$. Then there exists c such that for any $t \geq 0$,*

$$\partial r(t) \leq c\, [\Delta^{(m)} r](\lfloor t/m \rfloor),$$

and the sequence $\{[\Delta^{(m)} r](k)\}_k$ is a subgeometric sequence.

Proof. Since ∂r is non-decreasing, for any integer q,

$$\partial r(qm - m) \leq \int_{qm-m}^{qm} \frac{\partial r(s)}{m} \, \mathrm{d}s = \frac{\Delta^{(m)} r(q)}{m} \leq \partial r(qm). \tag{A.8}$$

Since $\partial r \in \Lambda_0$, there exists c such that for all t, $\partial r(t) \leq c\ \partial r(\lfloor t/m \rfloor m - m)$ thus yielding the first assertion. The second one is a consequence of the inequalities (A.8). □

Proof of Theorem 3.2. Let P^m be the irreducible skeleton. We can assume without loss of generality that $\Psi_1 \circ r_* \in \Lambda_0$, $\Psi_1 \circ r_* \geq 1$ and $\Psi_2 \circ f_* \geq 1$, and we do so.

We first prove that there exists c such that for any $t \geq 0$, $x \in \mathsf{X}$, there exists k and

$$\Psi_1 \circ r_*(t)\ \| P^t(x, \cdot) - \pi(\cdot) \|_{\Psi_2 \circ f_*} \leq c\ \Psi_1 \circ r_*(km)\ \| P^{km}(x, \cdot) - \pi(\cdot) \|_{\Psi_2 \circ f_*}. \tag{A.9}$$

Write $t = km + u$ for some $0 \leq u < m$ and a non-negative integer k. Since $\Psi_1 \circ r_* \in \Lambda_0$ and is a non-decreasing rate function, $\Psi_1 \circ r_*(km + u) \leq \Psi_1 \circ r_*(km)\ \Psi_1 \circ r_*(m)$. Furthermore, if $|g| \leq \Psi_2 \circ f_*$, upon noting that Ψ_2 and ϕ are non-decreasing concave functions

$$P^u |g| \leq P^u (\Psi_2 \circ \phi \circ V) \leq \Psi_2 \circ \phi \left(P^u V \right) \leq \Psi_2 \circ \phi \,(V + bm) \leq c\ \Psi_2(f_*),$$

where we used that by (3.1), $P^u V \leq V + bu$. This yields (A.9).

We now prove that the skeleton P^m is aperiodic and possesses an accessible and petite set A such that

$$\sup_A \mathbb{E}_x \left[\sum_{j=0}^{T_{m,A}-1} \Psi_1 \circ r_*(jm) \ \Psi_2 \circ f_*(X_{jm}) \right] < \infty. \tag{A.10}$$

Set $A = \{V \leq n\}$ for some n large enough: by Proposition 4.4 A is accessible and petite for the process and $\inf_{t \geq t_0} \inf_{x \in A} P^t(x, \cdot) \geq \nu(\cdot)$ for some t_0 and a maximal irreducibility measure ν ([22, Proposition 6.1] and [20, Proposition 3.2(ii)]). This implies that A and P^m have the desired properties. (A.10) is now a consequence of Proposition 4.8 and the inequality (3.3).

By using (A.9), the proofs of (3.4) to (3.7) are based on results on discrete-time Markov chains: Eq. (3.4) results from [30, Theorem 4.1, Eq(36)] while (3.5) is established in the proof of [30, Theorem 4.1]. (3.6) is a consequence of [30, Theorem 4.2].

By Lemma A.1 and the inequality (3.1), there exists $c < \infty$ such that

$$\begin{aligned} &\partial[\Psi_1 \circ r_*(t)] \, \|P^t(x, \cdot) - \pi(\cdot)\|_{\Psi_2 \circ f_*} \\ &\quad \leq c \ \Delta^{(m)}[\Psi_1 \circ r_*](\lfloor t/m \rfloor) \, \|P^{\lfloor t/m \rfloor m}(x, \cdot) - \pi(\cdot)\|_{\Psi_2 \circ f_*}. \end{aligned}$$

Since the sequence $\{\Delta^{(m)}[\Psi_1 \circ r_*](k)\}_k$ is equivalent to a sequence in the class Λ_0 defined e.g. in [30], (3.7) now follows from [30, Theorem 4.3]. $\square$

Proof of Theorem 3.3. (i) We first prove that $P^m W \leq W - \phi \circ W + b' \mathbf{1}_C$. This a consequence of Proposition 4.8 and Theorem 14.2.3(ii) in Meyn–Tweedie [21]. Indeed, since $\sup_C V < \infty$, indeed the condition on P^m shows that $\sup_{x \in C} \mathbb{E}_x \left[\sum_{k=0}^{T_{m,C}-1} \phi \circ V(X_{km}) \right] < \infty$. Define $\sigma_{m,C} = \inf\{k \geq 0, X_{mk} \in C\}$ and set $W(x) = \mathbb{E}_x \left[\sum_{k=0}^{\sigma_{m,C}} \phi \circ V(X_{km}) \right]$. Then the function W satisfies the conditions (see [21, Chapter 14]). As discussed in the proof of Theorem 3.2, for all $n \geq n_*$ the level sets $\{V \leq n\}$ are accessible and petite for the skeleton chain P^m. Choose then $\tilde{C} = \{V \leq n_* \vee \sup_C V\}$. (ii) The Moderate deviation principle (or MDP) comes from a decomposition into blocks and a return to the discrete-time case. Assume that $m = 1$ which can be done w.l.g. In fact, by (i), the Markov chain $(\Xi_k := X_{[k,k+1[})_{k \in \mathbb{N}}$ with probability transition Q is subgeometrically ergodic with the invariant probability measure $\tilde{\pi} = \mathbb{P}_\pi|_{\mathcal{F}_1}$ and satisfies **A1**–**A2** in the terminology of Douc–Guillin–Moulines [6]. Then, we may write

$$\begin{aligned} S_t^\epsilon &= \frac{1}{\sqrt{\epsilon} h(\epsilon)} \int_0^t g(X_{s/\epsilon}) \mathrm{d}s = \frac{\sqrt{\epsilon}}{h(\epsilon)} \sum_{k=0}^{\lfloor t/\epsilon \rfloor - 1} \int_k^{k+1} g(X_s) \mathrm{d}s + \frac{\sqrt{\epsilon}}{h(\epsilon)} \int_{\lfloor t/\epsilon \rfloor}^{t/\epsilon} g(X_s) \mathrm{d}s \\ &= \frac{\sqrt{\epsilon}}{h(\epsilon)} \sum_{k=0}^{\lfloor t/\epsilon \rfloor - 1} G(\Xi_k) + \frac{\sqrt{\epsilon}}{h(\epsilon)} \int_{\lfloor t/\epsilon \rfloor}^{t/\epsilon} g(X_s) \mathrm{d}s \end{aligned}$$

where G is obviously a bounded mapping with values in $\mathbb{R}^n$. By the boundedness of g, it is easy to see that the second term is exponentially negligible in the sense of moderate deviations, and thus S_t^ϵ and $\frac{\sqrt{\epsilon}}{h(\epsilon)} \sum_{k=0}^{\lfloor t/\epsilon \rfloor - 1} G(\Xi_k)$ are exponentially equivalent, and share the same MDP.

Note now that by Theorem 7 of Douc–Guillin–Moulines [6], under the subgeometric ergodicity of (Ξ_k) and the condition on the speed, $\frac{\sqrt{\epsilon}}{h(\epsilon)} \sum_{k=0}^{\lfloor t/\epsilon \rfloor - 1} G(\Xi_k)$ satisfies a MDP with speed $\frac{1}{h^2(\epsilon)}$ and rate function

$$\tilde{I}_g^h(\gamma) = \begin{cases} \dfrac{1}{2}\displaystyle\int_0^1 \sup_{\zeta\in\mathbb{R}^d} \left\{ \langle \dot\gamma(t), \zeta\rangle - \dfrac{1}{2}\tilde\sigma^2(\langle G, \zeta\rangle) \right\} \mathrm{d}t & \text{if } \mathrm{d}\gamma(t) = \dot\gamma(t)\mathrm{d}t, \gamma(0) = 0, \\ +\infty & \text{else,} \end{cases}$$

where $\tilde\sigma^2(\langle G, \zeta\rangle) = \lim_{n\to\infty} \frac{1}{n}\mathbb{E}_\pi\left(\sum_{k=0}^{n-1} G(\Xi_k)\right)^2$.

On the other hand, by the subexponential ergodicity, the boundedness of g and $\mathbb{E}_\pi\langle g, \zeta\rangle = 0$, we have that $\int_0^\infty (P^t\langle g, \zeta\rangle - \pi(\langle g, \zeta\rangle))\mathrm{d}t$ is absolutely convergent in $L^1(\pi)$. Thus $\tilde{I}_g^h = I_g^h$ as

$$\begin{aligned} \tilde\sigma^2(\langle G, \zeta\rangle) &= \lim_{n\to\infty} \frac{1}{n}\mathbb{E}_\pi\left(\int_0^n g(X_s)\mathrm{d}s\right)^2 = \lim_{n\to\infty}\frac{2}{n}\mathbb{E}_\pi\left(\int_0^n \mathrm{d}s \int_0^s \langle g, \zeta\rangle P^u\langle g, \zeta\rangle \mathrm{d}u\right) \\ &= 2\int_{\mathsf{X}} \langle g, \zeta\rangle \int_0^\infty P^u\langle g, \zeta\rangle \mathrm{d}u \,\mathrm{d}\pi = \sigma^2(\langle g, \zeta\rangle). \quad \square \end{aligned}$$

Proof of Theorem 3.4. (i) Since $V \in \mathcal{D}(\mathcal{A})$, there exists an increasing sequence $T_n \uparrow \infty$ of $\mathcal{F}_t$-stopping times such that for any n, $t \mapsto V(X_{t\wedge T_n}) - V(X_0) - \int_0^{t\wedge T_n} \mathcal{A}V(X_s)\mathrm{d}s$ is a $\mathbb{P}_x$-martingale. Set $T_{m,n} = \inf\{s \geq 0, |X_s| \geq m\} \wedge T_n$. There exists a constant $c < \infty$ such that on the set $\{s \leq T_{m,n}\}$, $V(X_s) + \phi\circ V(X_s) + |\mathcal{A}V(X_s)| \leq c$. This allows us to write

$$\begin{aligned} &\mathbb{E}_x[V(X_{t\wedge T_{m,n}})] + \mathbb{E}_x\left[\int_0^{t\wedge T_{m,n}} \phi\circ V(X_s)\mathrm{d}s\right] \\ &\quad = V(x) + \mathbb{E}_x\left[\int_0^{t\wedge T_{m,n}} [\mathcal{A}V(X_s) + \phi\circ V(X_s)]\mathrm{d}s\right] \leq V(x) + b\mathbb{E}_x\left(\int_0^t \mathbf{1}_C(X_s)\mathrm{d}s\right). \end{aligned}$$

The previous inequality is ensured by the monotone convergence theorem, that $\mathbb{E}_x\left[\int_0^t \phi\circ V(X_s)\mathrm{d}s\right]$ is finite. The proof is now completed by noting that

$$\begin{aligned} \mathbb{E}_x[V(X_t)] &= \mathbb{E}_x(\liminf_{n,m} V(X_{t\wedge T_{m,n}})) \leq \liminf_{n,m} \mathbb{E}_x(V(X_{t\wedge T_{m,n}})) \\ &\leq \liminf_{n,m}\left\{V(x) - \mathbb{E}_x\left[\int_0^{t\wedge T_{m,n}} \phi\circ V(X_s)\mathrm{d}s\right] + b\,\mathbb{E}_x\left[\int_0^{t\wedge T_{m,n}} \mathbf{1}_C(X_s)\mathrm{d}s\right]\right\} \\ &= V(x) - \mathbb{E}_x\left[\int_0^t \phi\circ V(X_s)\mathrm{d}s\right] + b\,\mathbb{E}_x\left[\int_0^t \mathbf{1}_C(X_s)\mathrm{d}s\right] \end{aligned}$$

where the last equality follows from monotone convergence.

(ii) (3.11) is trivial on C. Let $g(x) = \mathcal{A}V(x) + \phi\circ V(x)$. By $\mathbf{D}(\mathbf{C}, \mathbf{V}, \phi, \mathbf{b})$ and the definition of $\mathcal{A}V$, there exists an increasing sequence of stopping times $\{T_n, n \geq 1\}$ such that for any stopping time τ, $\mathbb{E}_x[\int_0^{\tau\wedge T_n}\{g(X_s) - b\mathbf{1}_C(X_s)\}\mathrm{d}s] \leq 0$. We prove that $g \leq 0$ on C^c. Let $x \notin C$ such that $g(x) \geq c > 0$ and set $\tau_x = \inf\{s \geq 0, g(X_s) - b\mathbf{1}_C(X_s) \leq 0.5c\}$. Since C is closed, g is continuous and $s \mapsto X_s$ is right continuous, $\mathbb{P}_x(\tau_x > 0) = 1$. Hence, $\mathbb{E}_x[\int_0^{\tau_x\wedge T_n}\{g(X_s) - b\mathbf{1}_C(X_s)\}\mathrm{d}s] \geq 0.5c\ \mathbb{E}_x[\tau_x \wedge T_n]$ which, by the monotone convergence theorem, is positive for n large enough. This is in contradiction with the assumptions and thus $g \leq 0$ on C^c. $\square$

References

[1] P. Billingsley, Convergence of probability measures, in: Wiley Series in Probability and Statistics: Probability and Statistics, 2nd ed., John Wiley & Sons Inc, New York, 1999, A Wiley-Interscience Publication.

[2] P. Cattiaux, I. Gentil, A. Guillin, Weak logarithmic sobolev inequalities and entropic convergence, Probab. Theory Related Fields 139 (3–4) (2007) 563–603.

[3] J.G. Dai, S.P. Meyn, Stability and convergence of moments for multiclass queueing networks via fluid limit models, IEEE Trans. Autom. Control 40 (11) (1995) 1889–1904.

[4] M.H.A. Davis, Markov models and optimization, in: Monographs on Statistics and Applied Probability, vol. 49, Chapman & Hall, London, 1993.

[5] R. Douc, G. Fort, E. Moulines, P. Soulier, Practical drift conditions for subgeometric rates of convergence, Ann. Appl. Probab. 14 (3) (2004) 1353–1377.

[6] R. Douc, A. Guillin, E Moulines, Bounds on regeneration times and limit theorems for subgeometric Markov chains, Ann. I.H.P. (Proba./Stat.) (2005) (in press).

[7] N. Down, S.P. Meyn, R.L. Tweedie, Exponential and uniform ergodicity of Markov processes, Ann. Probab. 23 (4) (1995) 1671–1691.

[8] S. Ethier, T. Kurtz, Markov processes, in: Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons Inc, New York, 1986, Characterization and convergence.

[9] G. Fort, Contrôle explicite d'ergodicité de chaîne de Markov: Applications à l'analyse de convergence de l'algorithme Monte-Carlo EM. Ph.D. thesis, Univ. P6, Paris, 2001.

[10] G. Fort, E. Moulines, V-subgeometric ergodicity for a Hastings-Metropolis algorithm, Statist. Probab. Lett. 49 (4) (2000) 401–410.

[11] G. Fort, G.O. Roberts, Subgeometric ergodicity of strong Markov processes, Ann. Appl. Probab. 15 (2) (2005) 1565–1589.

[12] H. Ganidis, B. Roynette, F. Simonot, Convergence rate of some semi-groups to their invariant probability, Stochastic Process. Appl. 79 (2) (1999) 243–263.

[13] R.K. Getoor (Ed.), Transience and recurrence of Markov processes, in: Lect. Notes Math., vol. 784, 1980, pp. 397–409.

[14] B. Goldys, B. Maslowski, Lower estimates of transition densities and bounds on exponential ergodicity for stochastic pde's, Ann. Probab. 34 (4) (2006).

[15] A. Guillin, Moderate deviations of inhomogeneous functionals of Markov processes and application to averaging, Stochastic Process. Appl. 92 (2) (2001) 287–313.

[16] R.Z. Has'minskii, Stochastic stability of differential equations. Sijthoff and Noordhoff, 1980.

[17] S. Jarner, G.O. Roberts, Polynomial convergence rates of Markov Chains, Ann. Appl. Probab. 12 (1) (2002) 224–247.

[18] M.A. Krasnosel'skii, Y.B. Rutickii, Convex functions and Orlicz spaces, Noordhoff, Groningen, 1961.

[19] M.N. Malyshkin, Subexponential estimates of the rate of convergence to the invariant measure for stochastic differential equations, Theory Probab. Appl. 45 (3) (2001) 466–479.

[20] S.P. Meyn, R.L. Tweedie, Generalized resolvents and Harris Recurrence of Markov processes, American Mathematical Society, Providence, RI, 1993, pp. 227–250.

[21] S.P. Meyn, R.L. Tweedie, Markov Chains and Stochastic Stability, Springer-Verlag London Ltd, London, 1993.

[22] S.P. Meyn, R.L. Tweedie, Stability of markovian processes II: Continuous-time processes and sampled chains, Adv. in Appl. Probab. 25 (1993) 487–517.

[23] S.P. Meyn, R.L. Tweedie, Stability of markovian processes III: Foster-Lyapunov criteria for continuous-time processes, Adv. in Appl. Probab. 25 (1993) 518–548.

[24] E. Nummelin, General Irreducible Markov Chains and Nonnegative Operators, Cambridge University Press, Cambridge, 1984.

[25] E. Nummelin, P. Tuominen, The rate of convergence in Orey's theorem for Harris recurrent Markov chains with applications to renewal theory, Stochastic Process. Appl. 15 (1983) 295–311.

[26] E. Pardoux, A.Yu. Veretennikov, On the Poisson equation and diffusion approximation. I,III, Ann. Probab. 29 (3) (2003) 1061–1085. 1111–1133, 2001.

[27] G.O. Roberts, R.L. Tweedie, Exponential convergence of Langevin diffusions and their discrete approximations, Bernoulli 2 (1996) 341–364.

[28] M. Röckner, F.Y. Wang, Weak Poincaré inequalities and L^2-convergence rates of Markov semigroups, J. Funct. Anal. 185 (2) (2001) 564–603.

[29] O. Stramer, R.L. Tweedie, Langevin-type models I: Diffusions with given stationary distributions, and their discretizations, Methodol. Comput. Appl. Probab. 1 (3) (1999) 283–306.
[30] P. Tuominen, R.L. Tweedie, Subgeometric rates of convergence of f-ergodic Markov Chains, Ann. Appl. Probab. 26 (3) (1994) 775–798.
[31] A.Yu. Veretennikov, On polynomial mixing bounds for stochastic differential equations, Stochastic Process. Appl. 70 (1) (1997) 115–127.
[32] L. Wu, Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems, Stochastic Process. Appl. 91 (2) (2001) 205–238.
[33] Z. Hou, Y. Liu, H. Zhang, Subgeometric rates of convergence for a class of continuous-time markov processes, J. Appl. Probab. 42 (2005) 698–712.